The Inference-Driven Model of Quantifier Focus

Eugene Jurij Dawydiak

Department of Psychology
University of Glasgow

Thesis submitted for the Degree of PhD to the Higher Degrees Committee of the Faculty of Social Sciences, University of Glasgow.

January, 2001

Abstract

Previous research (e.g. Moxey & Sanford, 1987) has demonstrated that positively and negatively quantified statements place attentional focus upon different subsets of a logical superset and are, thus, associated with different patterns of subsequent anaphoric reference. Typically, positive quantifiers (e.g. *many (x)*) are associated with anaphoric reference to the subset of entities for whom the sentence predicate is true (the reference set or *refset*) while negative quantifiers (e.g. *not many (x)*) are associated with reference to the subset for whom the predicate is false (the complement set or *compset*). The primary objective of this thesis is to provide an empirical evaluation of a model (the Inference Model) which has been advanced to explain the occurrence of complement set reference. The experiments reported in this thesis address three central issues associated with the Inference Model: the relationship between denial and complement set reference; the impact of connective on complement set reference; and, the time course over which complement set focus is established. The empirical chapters contained in Section B present the results from four off-line sentence completion studies. Experiments 1 and 2 investigated the relative effects of denial and downwards monotonicity on complement set reference in the context of both free continuation (Experiment 1) and forced reference (Experiment 2) tasks. Consistent with the Inference Model, the results of both studies indicated that denial was the major determinant of complement set reference. Experiment 3 represented a slight digression from the main direction of the thesis and was conducted in order to test an explanation for an unexpected observation made in the first two experiments. Experiment 4 investigated the sensitivity of set reference pattern to the inferential constraints imposed by different connectives. The results indicated that the pattern of set reference was largely insensitive to connective and were interpreted as being consistent with focus having been established prior to the connective being encountered. Section C reports the results from three on-line studies which attempt to identify the time course over which complement set focus develops. This question is important as the Inference Model predicts that compset focus should be established less immediately than refset focus. Experiments 5 and 6 measured reading times for an anaphoric reference sentence

which appeared immediately after a quantified statement. The results demonstrated no systematic asymmetry in the time taken to resolve compset and refset anaphors and were interpreted as being consistent with focus for both subsets having been established during comprehension of the quantified sentence. Experiment 7 attempted to index the time course of set focus effects using the more sensitive on-line measure of eye-movements. The results were generally consistent with the Inference Model in that they suggested that compset focus was established later than refset focus. The overall pattern of results presented here can be interpreted as being broadly consistent with both the Inference Model and a less complex mechanism based upon the general account of negative processing suggested by Clark (1976). The results are not, however, consistent with an alternative account of complement set reference proposed by Kibble (1997).

Acknowledgements

I wish to thank my supervisor Tony Sanford for his constant encouragement, support and extreme patience throughout the course of this research. I also wish to thank my second supervisor Linda Moxey for her support and guidance during this period.

Thanks to my friends, Andrew, Annie, Asifa, Jason, Margaret, Paul and Siobhan, for making the whole experience fun.

Thanks to Julian for enduring the torment of proof-reading and thanks to Christine for her endless support and encouragement.

For Jurij and Elizabeth.

Declaration

I declare that this thesis is my own work carried out under normal terms of supervision (but see footnote for Chapter 7 on page 172).

Contents

List of Figures

Chapter Two

Chapter Three

Chapter Four

Chapter Five

Chapter Six

Chapter Seven

Chapter Eight

List of Tables

Chapter One

Natural Language Quantifiers and Psychological Function

1.1 Introduction

A considerable volume of work within the field of quantifier research has focused upon issues associated with both the scope of quantification (e.g. Catlin & Micham, 1975; Chien, 1994; Kemtes & Kemper, 1999; Kurtzman & MacDonald, 1993; Micham, Catlin, VanDerveer, & Loveland, 1980) and the dynamics of quantificational reasoning (e.g. Ekberg & Lopes, 1979, 1980; Johnson-Laird, 1983; Johnson-Laird & Byrne, 1989, 1991; Newstead, 1995; Revlis, Lipkin, & Hayes, 1971; Rosenthal, 1980). This thesis is concerned with a somewhat different aspect of quantification. This work concerns the patterns of anaphoric reference which are associated with natural language quantifiers and represents a continuation of a line of research initiated by Moxey (1986). In more specific terms, the thesis examines the phenomenon of *complement set* (or *compset*) reference. The general aim of this work is to empirically evaluate, and further elaborate, a theory which has been advanced by Sanford, Moxey and Paterson (1996; Moxey & Sanford, 2000a) to explain the occurrence of complement set reference. The main body of this thesis is separated into three sections. Section A brings together the relevant psychological literature in order to provide a comprehensive description of the phenomenon and allow an initial evaluation of the explanatory mechanisms which have been advanced to account for it. The remaining sections relate the results from a series of experiments which investigate the mechanism which underlies compset reference in the context of both language production and comprehension. Section B reports the results from four off-line sentence completion studies and section C, the results from three on-line reading experiments.

The aim of this first chapter is to provide a general conceptual context for the central issues addressed by the thesis. It begins by providing a formal definition for the term *quantifier*. It then outlines the traditional approach to the understanding of quantifier function and identifies several critical problems which are associated with this. Finally, the chapter relates evidence which indicates that quantifiers serve a more complex psychological function than is suggested by the traditional approach.

1.2 Quantifiers: A Formal Definition

Within the classification framework of Generalised Quantifier Theory (GQT) (Barwise & Cooper, 1981; Keenan & Stavi, 1986), a natural language quantifier is described in terms of two components: a determiner, which can be characterised as the quantity expression (e.g. *few, a few, many, not quite all, ten*) and a singular or plural set expression. In formal terms, this can be expressed as $Qx(x)$, where Qx is the determiner and (x) is the set expression. It should be noted, however, that certain authors have used the term *quantifier* to refer to the determiner in isolation. For the sake of clarity, the Barwise and Cooper terminology will be maintained throughout the thesis. Quantifiers therefore correspond to quantified noun-phrases rather than simply the determiner itself. For proper names, the components are implicit in the term as a set composed of one entity is specified. Where a plural set is introduced, the quantifier can be interpreted as specifying a relationship between this superset and its logical subsets. For example, given sentence (1), the following sets can be logically derived: set A: the superset of children, set B: the subset of children for whom the predicate is true (i.e. the children who were smiling) and set C, the complement of B, the subset of children for whom the predicate is false (i.e. those who were not smiling).

(1) Some of the children were smiling.

Determiners can therefore be seen as serving to *partition* the set expression in terms of some property which is true of one subset, but not of the other.

1.3 The Traditional Approach to Quantifier Function: Quantifiers as Descriptions of Amount

At the outset, two observations would appear to be pertinent to the understanding of quantifier function. First, on the surface, quantifiers as a group appear to describe poorly defined amounts or quantities. Second, they, along with other related linguistic expressions (such as frequency adverbs and probability statements), are members of a class of linguistic expression which can be understood in terms of scales (e.g. Jarvella & Lundquist, 1994; Lundquist & Jarvella, 1994; Moxey & Sanford, 1993a). Thus, quantifiers seem to express vague amounts which are differentiable with respect to scalar position (e.g. *many (x)* is more than *not many (x)*, and *almost none (x)* is less than *nearly all (x)*). Given this initial information, it would seem reasonable to conclude that quantifiers simply function as a linguistic device for the communication of vague quantity information. It is evident, however, that in order for quantifiers to function in this respect, two requirements would necessarily have to be satisfied. Individual communicators would have to maintain relatively stable scalar interpretations of the quantifiers (necessary for the meaningful representation of quantity) and these interpretations would have to be consistent between communicators (necessary for the meaningful communication of quantity). This section assesses the extent to which quantifiers fulfil these requirements.

1.3.1 Quantifier scalar consistency within and between individuals

A considerable amount of research effort has been directed towards attempting to map quantity expressions on to numerical values in order to determine the relative rank order of the amounts expressed by different quantifiers. Such studies typically require the participant to generate estimates of the proportion or number denoted by each member of a set of different quantifiers. One specific technique used is that of magnitude estimation. Here the participant is initially asked to provide an estimate for a given expression and is then required to generate further estimates for additional expressions relative to this

anchor point. The results from studies which have used this methodology (e.g. Bass, Cascio & O'Connor, 1974; Goodwin, Thomas & Hartley, 1977; Hammerton, 1976; Pohl, 1981; Schriesheim & Schriesheim, 1974, 1978) generally indicate that it is possible for participants to differentiate small sets of quantifiers with respect to the numbers or proportions they denote. Moreover, some studies have indicated that, for individual participants, such scalar mappings remain relatively stable over time (e.g. Beyth-Marom, 1982; Budescu & Wallsten, 1985).

The evidence related above would seem to suggest that quantifiers can, under certain circumstances, be used to meaningfully represent quantity information. However, there are problems associated with such an interpretation of the evidence. First, as Moxey and Sanford (1993a, 2000b) remark, scaling studies typically report that the estimates generated by participants in response to different quantifiers show a considerable degree of overlap. As a consequence of this, the studies which do report relatively discrete quantifier to amount mappings for individuals tend to be those which employ small sets of, maximally discriminable, quantifiers as test items.

A second source of problems for scaling studies derives from the nature of the task employed itself. In many of these studies, participants are required to generate an estimate for each of a series of quantifiers and this procedure can give rise to, what have been termed, contrast effects. These effects have been demonstrated by a considerable volume of work (e.g. Chase, 1969; Poulton, 1973, 1989) which indicates that later estimates of amount are likely to be influenced by those estimates which have already been made by the participant. That is, initial estimates of magnitude bias those made later in the test session. This observation is consistent with findings from more general work in the field of heuristics (e.g. Tversky & Kahneman, 1974), concerning the biasing effect of adjusting estimates from an established anchor point.

The existence of contrast effects raises the possibility that the quantifier to amount mappings obtained in some work may, at least in part, be attributable to weaknesses in the

methodological approach adopted. This possibility was directly addressed by Moxey and Sanford (1993b). In order to avoid contrast effects, Moxey and Sanford asked each of 450 participants to provide a single percentage estimate of the amount denoted by one of ten different quantifiers. The results indicated that there was no reliable difference between the estimates generated in response to *very few (x), few (x), a few (x), not many (x)* and *only a few (x).* Thus, Moxey and Sanford concluded that in terms of the quantity information communicated, these low-ranking quantifiers effectively denoted a similar amount.

The general lack of differentiation in the amount perceived to be denoted by different quantifiers presents a clear difficulty for the idea that quantifiers function to allow the communication of meaningful quantity information. The inability of quantifiers to function in this respect is further demonstrated by the findings from several studies which indicate that the scalar mappings associated with specific quantifiers are highly variable between individuals (e.g. Beyth-Marom, 1982; Budescu & Wallsten, 1985; Johnson, 1973).

1.3.2 Quantifiers and membership functions

The evidence reviewed above would appear to be inconsistent with the view that quantifiers simply map on to specific quantities. An alternative, and more complex, approach to the study of vague statements of amount and frequency makes use of the concept of fuzzy-set membership functions. This approach involves the expression of category membership in terms of graded acceptability, rather than in terms of an absolute mapping. Here, category membership judgements are expressed on a scale of 0 through 1, where "0" represents items which are certainly not within a given category and "1" represents items which certainly do fall within the category (Moxey & Sanford, 1993a). Wallsten, Budescu, Rapoport, Zwick and Forsyth (1986) claim that this approach can be applied to understanding the meaning of natural language expressions of probability and frequency. Thus, they argue that such expressions do not map onto specific numerical

equivalents but, rather, map onto ranges of acceptable values. Wallsten et al (1986; Rapoport, Wallsten & Cox, 1987), conducted an empirical investigation of the membership functions associated with different probability terms. The results generally appeared to indicate that these remained relatively constant for specific individuals across different tasks. However, as with the less sophisticated investigations described above, the results also revealed a lack of consistency between individuals. This finding is, again, not consistent with the view that such terms function as a medium for the communication of meaningful scalar information.

1.3.3 Other sources of variation in the amount denoted: The influence of context

An additional problem for the idea that quantifiers simply translate into vague statements of amount arises from several strands of evidence which suggest that the amount perceived to be denoted by a given quantifier is sensitive to the context in which it is used. At an intuitive level, this point can be illustrated with reference to Parducci's (1968) observation that while *almost never* may be considered an appropriate statement of frequency for "missing 5% of lectures", *often* may be more appropriate in the case of a "5% failure rate for contraceptives". Thus, within different contexts, both high and low ranking frequency adverbs would appear to be appropriate descriptions of the same actual frequency.

There exists a considerable volume of empirical work which demonstrates the influence of context associated baserate expectation upon frequency estimates. For example, Newstead and Collis (1987) reported the results from four experiments which investigated the effect of context upon number estimates generated in response to frequency adverbs (e.g. *often* and *seldom*). A central finding was that the proportion estimates produced were sensitive to the expected frequency of a given event. Thus, *often* was perceived to signify a higher frequency in the context of trips to the cinema than it did within the context of visits to the USA. Additionally, Newstead and Collis also reported that proportion estimates were

sensitive to emotional context. Here, frequency adverbs were interpreted as communicating higher proportions in the context of feeling *happy* (assumed to be a generally high frequency occurrence) than they were for feeling *suicidal* (assumed to be a generally low frequency occurrence). The general pattern evident in this, and other research (e.g. Pepper & Prytulak, 1974; Wallsten, Fillbaum, & Cox, 1986), is that frequency expressions tend to be perceived as denoting higher frequencies in the context of higher baserate expectations.

Context has also been demonstrated to have an impact upon the numerical values assigned to quantifiers. Moxey and Sanford (1993b) demonstrated an effect of baserate expectation in a study which investigated the influence of three situational contexts upon the proportions generated in response to ten different quantifiers. Each of the three contexts employed was associated with a different baserate quantity expectation (low, medium or high). Their results indicated that the quantity estimates generated by participants for the higher ranking quantifiers in the study were sensitive to baserate expectation. In line with previous research, these quantifiers were associated with increasingly larger proportions as baserate expectation increased.

The values assigned to quantifiers have been shown to be sensitive to a range of other contextual elements. Hörmann (1983) found that the numbers assigned to *einige*, *mehrere* and *ein paar* (*some*, *several* and *a few*) were sensitive to object size, type and location. Newstead, Pollard and Reizbos (1987) found that the proportion of a set judged by participants to be associated with a given quantifier was sensitive to the size of that set. More recently, Newstead and Coventry (2000) investigated the effect of set member size and functional relationship upon the perceived appropriateness of different quantifiers used to describe the set. They found that both the size of set members (in the experiments, these were balls) and the functional relationship of these with their environment (a bowl) had an impact upon the judged appropriateness of quantifiers used to describe the set. From these results, Newstead and Coventry concluded that quantifiers derive most of their meaning from the context of their usage.

1.4 The Psychological Effects Associated with Quantifiers

The evidence considered above was primarily concerned with the problems which exist for the view that quantifiers simply serve to facilitate the communication of vague quantity information. This section considers two strands of evidence which suggest that quantifiers may be implicated in more complex psychological function.

1.4.1 Quantifiers and attribution

Attribution Theory, a field of research originated by Heider (1958), deals with the ways in which causes are assigned to behaviour. Within this area of research, there exist a number of distinct theoretical perspectives (e.g. Bem, 1962; Jones & Davis, 1965; Schachter, 1964) which emphasise the importance of different factors in the attributional process. One factor which has been identified as an important component of attribution is the presence of relative frequency and proportion information (Kelly, 1967, 1973). The impact of these types of information can be illustrated with reference to the information related in (2).

(2) Jane shouted at Roger.

Here, two immediately plausible reasons for the action described are that it is caused by a property of Jane (e.g. she has a short temper) or that it is caused by a property of Roger (e.g. he is an annoying person). In such a situation frequency and proportion information may be important in the assigning of a cause to the action. For example, if Jane is known to frequently perform the action, this may suggest that the reason lies with her; however, if a relatively large proportion of other people are also known to perform the action, this may suggest that cause is attributable to Roger.

Barton and Sanford (1990) suggest that quantifiers and frequency adverbs may be implicated in the control of attributions. They conducted an experiment in which

participants were presented with a set of two-sentence materials (of the type shown in Table 1.1) and asked to make an attribution judgement. These materials were designed to relate information concerning the frequency of occurrence of some given action for both an individual and the general population. They found that while certain low ranking frequency adverbs and quantifiers (*rarely, only occasionally, only a few* and *few*) had an impact upon attributions, others which signalled a similarly low frequency (*occasionally* and *a few*) did not.

Table 1.1 A typical two sentence material from the Barton and Sanford (1990) study

John * enjoys walking his dog.
** other people enjoy walking their dogs.
Where: * was one of: *rarely, only occasionally, occasionally* or blank ; and ** was one of: *a few, only a few* or *few*.

Barton and Sanford argued that, since the main test expressions all communicated a similar frequency, the observed differences in attributional pattern could not be explained in terms of the frequency information alone. They concluded that the effect was confined to *rarely, only occasionally, only a few* and *few* because these expressions focus the readers attention upon causes and, consequently impact upon attributions. This finding again suggests that quantifiers do not function solely to mediate the communication of quantity information.

1.4.2 The reference patterns associated with scalarity

Jarvella and Lundquist (1994; Jarvella, Lundquist & Hyönä, 1995; Lundquist & Jarvella, 1994) have pursued a line of research which directly links scalarity with referential focus. This research, which is based upon a linguistic theory developed by Ducrot and Anscombre (Anscombre, 1989, 1991; Anscombre & Ducrot, 1986; Ducrot, 1988, 1995), is underpinned by the observation that certain scalar expressions (e.g. *almost*) appear to

be ascending and are interpreted as indicating a comparatively large amount or proportion, while others (e.g. *only*) seem to be descending and suggest a small amount (Jarvella, Lundquist & Hyönä, 1995). These properties of ascending and descending scalarity are also held to interact with existing world knowledge (relating to the meaning of high or low numbers within a given context) and can, as a consequence, impact upon the acceptability of subsequent anaphoric reference. This effect can be illustrated with reference to texts (2) - (5) (from Jarvella, Lundquist & Hyönä, 1995, p 6) below.

(2) John Smith got almost 500 votes in the first round. He is likely to win the election. √

(3) John Smith got almost 500 votes in the first round. He is likely to lose the election. ×

(4) John Smith got only 500 votes in the first round. He is likely to lose the election. √

(5) John Smith got only 500 votes in the first round. He is likely to win the election. ×

In (2) and (3), the ascending scalarity of *almost* indicates that a large amount of votes were gained relative to some unspecified benchmark. This information then interacts with the world knowledge that "the more votes someone gets, the more likely that person is to win". Thus, (2) represents an apparently acceptable sequence of text since it allows *he* to be interpreted as being co-referential with *John Smith*, while (3) does not appear to be acceptable (since *he* does not appear to be interpretable as being co-referential with *John Smith*). Similarly, for (4) and (5), the descending scalarity of *only* (which specifies a low amount relative to some benchmark) interacts with the world knowledge that "the less votes someone gets, the more likely that person is to lose." Thus rendering (4) an

acceptable statement (*he* can again be interpreted as referring to *John Smith*) and (5) unacceptable.

The general principle described above can be also be extended to cases where the anaphor is a full NP rather than a pronoun as illustrated by (6) and (7) (from: Lundquist & Jarvella, 1994, p 39). In (6), *Nichols* and *the Irish lad* appear to be co-referential, while in (7), for the text to be interpretable, *the Irish lad* must be understood as referring to some other character who is left unspecified in the text.

(6) Nichols obtained *almost* 129 points. The red-haired Irish lad seemed to be leading.

(7) Nichols obtained *only* 129 points. The red-haired Irish lad seemed to be leading.

Jarvella and Lundquist (1994; Lundquist & Jarvella, 1994) substantiated these claims in a series of on- and off-line experiments which utilised both Danish and Finnish scalar expressions. For example, the studies reported by Jarvella and Lundquist (1994) demonstrated the effect of scalar direction on referential access. These used a series of experimental items of the type (translated into English) shown in (8) below. These began with a context setting description of a competitive event and the introduction of two named characters. Following this, a further two sentences were included. The first of these related information concerning the competitive performance of one of the previously introduced characters and included one of two scalar expressions (either the descending quantifier *only a few (x)* or the ascending quantifier *almost all (x)*). The second sentence began with a noun-phrase (*the tall blond Århus-girl* or *the 20-year-old office girl*) which acted as an anaphoric reference to one of the previously mentioned characters. This sentence also included the scalar information that the person in question was either *behind* or *ahead.*

The central prediction made by Jarvella and Lundquist was that the noun-phrase in s2 would only be interpreted as being co-referential with the character named in s1 when the scalar expressions ran in the same direction. In terms of the example given in (8), *the tall blond Århus-girl* would only be interpreted as referring to *Karina Madsen* when the scalar expressions were either *almost all* and *ahead* or *only a few* and *behind.* Jarvella and Lundquist measured referential focus in two ways. In the off-line study participants were required to indicate which characters they thought were winning or losing. In the on-line experiment, participants responded *yes* or *no* to questions of the form *Did N win?* and *Did N lose?* The results of both studies supported the experimental prediction.

(8) The Beauty Competition

The year's big beauty competition took place in the Circus Theatre in Copenhagen. Eighteen models participated, chosen from various provincial cities. As journalists and other media people had suspected, Karina Madsen and Louise Fernholm stood out as the big favourites.

s1 In the round with bathing suits, Karina Madsen received high marks from *only a few/almost all* the judges.

s2 The tall blond Århus-girl was clearly *behind/ahead.*

These findings are important because they suggest that quantifier scalar direction is associated with specific types of inferential activity and, moreover, that this inferential activity has a direct impact upon the pattern of subsequent anaphoric reference.

1.5 Summary and Conclusion

The majority of this chapter was concerned with an evaluation of the contention that the function of quantifiers is simply to facilitate the communication of vague quantity information. It was argued that in order to mediate this function, individual quantifiers would have to map onto relatively stable, and differentiable, amount equivalents both within and between individuals. A substantial amount of evidence was then reviewed which suggested that quantifiers do not fulfil these requirements.

The chapter then considered two lines of research which suggest that quantifiers may be associated with more complex psychological function. First, there is some evidence to suggest that quantifiers may impact upon the attributional process. Secondly, there is strong empirical evidence which links quantifier scalar direction with both inferential activity and reference pattern. This latter finding is of considerable relevance to the phenomenon of complement set reference. The nature of this phenomenon is examined in the forthcoming section of this thesis.

Section A

A Review of the Phenomenon

This section provides a detailed description of the phenomenon of complement set reference and examines the mechanisms which have been advanced to explain its occurrence.

Chapter Two

The Focus Effects Associated With Natural Language Quantifiers

2.1 Introduction

The principal purpose of this chapter is to provide a description of the phenomenon of variable set focus which is associated with natural language quantifiers. The chapter begins by providing a conceptual framework for the classification of quantifiers in terms of monotonicity and polarity. It then reviews the empirical evidence relating to the linguistic set-focusing effects associated with quantifiers and considers the conditions which license these effects. Finally, the chapter discusses the alternative interpretations of the data which have been suggested by researchers.

2.2 Quantifier Classification

This section provides a conceptual framework for the classification of quantifiers in terms of monotonicity and quantified sentences in terms of polarity. Additionally, it introduces a series of natural language tests which act as diagnostics of these properties.

2.2.1 Quantifier monotonicity

With GQT, quantifiers can be characterised in terms of the type of relationship they signal between the subset and the superset. **Monotone increasing** (M↑) quantifiers signal a quantity assertion about a subset which is also necessarily true of the superset OR a quantity assertion about the superset which is not necessarily true of a subset. So, if the quantifier can be interpreted as asserting that an upwardly unbounded amount is true of a subset (i.e. *x or more than x*), then it follows that this must also be true of the superset

(since it permits the possibility that the superset has an equal or greater number of elements). Additionally, it also follows that a uniquely identified subset of the superset (as in *the x of the...*) will be M↑ (ie, if specified elements *x* and *y* constitute a subset of the superset, it follows that these specified elements must be present in the superset). **Monotone decreasing** (M↓) quantifiers signal inferences in the opposite direction. Here, what is true for the subset, is not necessarily true of the superset but what is true of the superset will be true of the *subset*. This will be true if the quantifier asserts a downwardly unbounded amount (i.e. *x or less than x*) is true of the superset (since it allows for the subset to have an equal or lesser number of elements). **Nonmonotonic** quantifiers do not signal inferences of this type. Here, the quantity expression is neither upwardly nor downwardly unbounded. This is the case for quantifiers where the quantity is precisely, but not uniquely, specified (as in *exactly two*) and also where the quantifier is an approximative (as in *about 2* or *roughly 2* and so on). Barwise and Cooper (1981) suggest that such nonmonotonic quantifiers exist as composites of their M↓ and M↑ counterparts. So, for example, *exactly two* can be modelled as a composite of *no more than two* (M↓) and *no less than two* (M↑).

2.2.2 Natural language tests of monotonicity

The type of set/subset relationship signalled by a quantifier's monotonicity can also be seen to have consequences for the apparent acceptability of the quantifier's usage in certain sentences. This linguistic feature has been employed by some investigators (e.g. Moxey & Sanford, 1993a) as the basis for a natural language diagnostic of quantifier monotonicity. The test involves the substitution of the determiner for the quantity expression *Qx* in sentences (1) and (2) below. The degree to which the resulting sentence forms an intuitively acceptable assertion can then be taken as an index of the quantifiers monotonicity.

(1) If *Qx* of the fans went to the concert early, then *Qx* of the fans went to the concert.

(2) If *Qx* of the fans went to the concert, then *Qx* of the fans went to the concert early.

If the determiner substituted for *Qx* yields a true statement for sentence (1) but not sentence (2), the quantifier can be held to be M↑. For example:

(3) If *more than 10%* of the fans went to the concert early, then *more than 10%* of the fans went to the concert. √

(4) If *more than 10%* of the fans went to the concert, then *more than 10%* of the fans went to the concert early. ×

It can be seen that, for (3) the right-hand *then...* assertion necessarily follows from the left-hand *if...* statement, while, for (4) this is obviously not the case. The apparent truth of (3), and falsity of (4), indicates that *more than 10%* is monotone increasing.

Similarly, if the quantifier yields a true statement for (2) but not (1), the quantifier is M↓. In (5) and (6) below, the pattern of true statements is reversed.

(5) If *less than 10%* of the fans went to the concert early, then *less than 10%* of the fans went to the concert. ×

(6) If *less than 10%* of the fans went to the concert, then *less than 10%* of the fans went to the concert early. √

Substitution of a nonmonotone quantifier for *Qx(x)* in the sentences leads to a true statement for neither (as in (7) and (8)).

(7) If *exactly two* of the fans went to the concert early, then *exactly two* of the fans went to the concert. ×

(8) If *exactly two* of the fans went to the concert, then *exactly two* of the fans went to the concert early. ×

Although some quantifiers are analytically M↓ or M↑ (such as *less than 10* or *more than 10*), others emerge as apparently M↓ or M↑ from the intuitive acceptability of their substitution in tests like those above. For example, when the quantifier *few* is substituted for *Q(x)* in the tests above, it yields an apparently unacceptable assertion in (9) but not (10); indicating that *few* is M↓. Similarly, substitution of *many* indicates the opposite pattern of acceptability ((11) is acceptable, while (12) seems unacceptable).

(9) If *few* of the fans went to the concert early, then *few* of the fans went to the concert. ×

(10) If *few* of the fans went to the concert, then *few* of the fans went to the concert early. √

(11) If *many* of the fans went to the concert early, then *many* of the fans went to the concert. √

(12) If *many* of the fans went to the concert, then *many* of the fans went to the concert early. ×

Table 2.1 provides a summary of quantifier monotonicity classification based on the apparent truth values obtained through quantifier substitution in the test sentences.

Table 2.1 Quantifier monotonicity classification in terms of the natural language test sentences

Analytically Monotone Decreasing	Intuitively Monotone Decreasing	Analytically Monotone Increasing	Intuitively Monotone Increasing	Undecideable/ nonmonotone
Less than x (x)	*Few x (x)*	*More than x (x)*	*A few x (x)*	*Not quite x (x)*
None (x)	*Not many x (x)*	*All (x)*	*Many x (x)*	*Almost x (x)*
At most x (x)		*At least x (x)*		*Nearly x (x)*
				Only a few (x)

2.2.3 Sentence polarity and tag questions

Sentences may be characterised in terms of polarity. This refers to whether the sentence can be interpreted as an affirmation or a negation. Klima (1964) observed that the presence of a negative particle licenses a different tag question in a declarative sentence from one that does not contain such a particle. In other words, positive polarity sentences combine grammatically with negative tag questions (such as *didn't he?*), as in (13), but not with positive tag questions (such as *did he?*), as in (14).

(13) John **did** go to the concert, **didn't he?** √

(14) John **did** go to the concert, **did he?** ×

For negative polarity sentences, the pattern is reversed. Sentences which assert a negation combine grammatically with positive tag questions, as in (15), but not with negative tag questions, as in (16).

(15) John **didn't** go to the concert, **did he?** √

(16) John **didn't** go to the concert, **didn't he?** ×

Intuitive judgements concerning the grammaticality of sentences combined with tag questions can therefore be used as a diagnostic of sentence polarity. This, in principle, provides an additional tool for classifying quantifiers. Given a simple declarative sentence of the form *Qx(x) went to the concert, did/didn't they?*, it is possible to substitute different quantifiers for *Qx(x)* and assess the polarity of the resulting sentence by way of the apparent grammaticality of the results. For example, the apparent acceptability of sentences (17) and (18), in contrast to sentences (19) and (20), would seem to indicate that *many (x)* is positive, while *not many (x)* is negative.

(17) Many of the politicians went to the concert, didn't they? √

(18) Not many of the politicians went to the concert, did they? √

(19) Many of the politicians went to the concert, did they? ×

(20) Not many of the politicians went to the concert, didn't they? ×

2.2.4 Downwards monotonicity and negative polarity items

Negative polarity items (NPIs), which include terms such as *anymore, give a damn* and *give a red cent*, are held to occur within the scope of negations (e.g. Van der Wouden, 1997) and can consequently be used as a diagnostic of a linguistically negative environment. This principle is illustrated by the acceptability of sentence (21), in contrast with the unacceptability of sentence (22).

(21) John doesn't like ice-cream anymore. √

(22) John likes ice-cream anymore. ×

A further natural language test for M↓ (and negative polarity) quantifiers derives from the claim that M↓ quantifiers license NPIs (e.g. Ladusaw, 1979; van der Wouden, 1997; Zwarts, 1996). (23) and (24) below illustrate that negatively quantified declarative

statements combine grammatically with negative polarity items while (25) and (26) show that this is not the case for the positively quantified counterparts.

(23) Not many/Few/None of the students take drugs anymore. √

(24) Not many/Few/None of the students give a damn about lectures. √

are acceptable, while (25) and (26) are not.

(25) Many/A few/All of the students take drugs anymore. ×

(26) Many/A few/All of the students give a damn about lectures. ×

2.2.5 Issues arising from the application of the tests

The tests described above provide, in principle, a method for the classification of quantifier monotonicity and the polarity of declarative sentences containing them. Table 2.2 below provides a summary of the resulting classification of quantifiers in terms of these tests. From this table, it is apparent that many quantifiers produce a consistent pattern of results in that quantifier downwards monotonicity is associated with intuitively judged acceptance of NPIs and the formation of negative sentences as judged by tag test results (as is demonstrated by (27)) while upwards monotonicity is associated with non-acceptance of NPIs and the formation of positive sentences.

(27) If *not many* of the fans went to the concert, then *not many* of the fans went to the concert early. √

Not many people do that anymore √

Not many people went to the party, did they? √

Table 2.2 Quantifier monotonicity and polarity in terms of natural language tests

Quantifier	Monotonicity	Negative Polarity Item	Tag Test
Not all (x)	M↓	yes	-ve
Not many (x)	M↓	yes	-ve
Few (x)	M↓	yes	-ve
Not quite X (x)	?	yes	-ve
No more than X (x)	M↓	yes	-ve
At most X (x)	M↓	yes?	+ve
Less than X (x)	M↓	yes?	+ve
Only a few (x)	?	yes	+ve?
Almost all (x)	?	no	+ve
Many (x)	M↑	no	+ve
A few (x)	M↑	no	+ve
Nearly X(x)	?	no	+ve
More than X (x)	M↑	no	+ve
At least X (x)	M↑	no	+ve
More than X (x)	M↑	no	+ve

Where:
Monotonicity refers to the acceptability, or otherwise, of the quantifiers substitution into the following sentences:
If *Qx* of the fans went to the concert early, then *Qx* of the fans went to the concert. (M↑).
If *Qx* of the fans went to the concert, then *Qx* of the fans went to the concert early. (M↓).
M↑ indicates monotone increasing; M↓ indicates monotone decreasing and "?" indicates nonmonotone or undecidable results.
Negative Polarity Item refers to the grammatical combination of a quantified statement with a negative polarity item. "?" indicates mixed or undecidable results, based on intuitive judgements.
Tag Test refers to a quantifier polarity judgement in terms of the acceptability of the quantified statements' combination with positive and negative tag questions. "?" indicates undecidable results.

There are, however, cases which do not fit this pattern. In particular, two main categories of atypical results are evident. First, there are some quantifiers which do not appear to be M↓ by the monotonicity sentence tests, but which do accept NPIs and form negative sentences. This is apparently the case for *not quite X (x)* , as is illustrated by the sentences in (28) and (29).

(28) If *not quite 50%* of the fans went to the concert, then *not quite 50%* of the fans went to the concert early. ×

Sentence (28) shows that *not quite 50%(x)* is not M↓ because the quantifier does not permit the possibility for the set of fans who went early to be empty (i.e. it is possible that none of the fans went early). However, as illustrated by (29), the quantifier does seem to both accept NPIs and form negative sentences.

(29) *Not quite 50%* of the people do that anymore √

Not quite 50% of the people went to the party, did they? √

A second category of atypical results concerns quantifiers which do seem to be M↓ and may accept NPIs, but form positive sentences as judged by tag tests. The quantifiers *at most X (x)* and *less than X (x)* exemplify this category, as is illustrated below.

(30) If *at most 50%* of the fans went to the concert, then *at most 50%* of the fans went to the concert early. √

If *less than 50%* of the fans went to the concert, then *less than 50%* of the fans went to the concert early. √

For (30), true statements are generated for both quantifiers, indicating that they are M↓. Additionally, the sentences in (31) seem to suggest that the quantifiers may accept NPIs. The sentences, however, do sound rather laboured and unnatural.

(31) *At most 50%* of the people do that anymore √?

Less than 50% of the people do that anymore √?

Sentences (32), however, indicate that these quantified statements do accept negative tag questions and therefore appear to be positive in polarity. This would seem to suggest that quantifier downwards monotonicity is not always associated with the generation of an unambiguously negative linguistic environment. Moreover, as judgements of negativity

(based upon acceptance of NPIs and positive tag questions) do not always appear to yield consistent results.

(32) *At most 50%* of the people went to the party, did they? ×

At most 50% of the people went to the party, didn't they? √

Less than 50% of the people went to the party, did they? ×

Less than 50% of the people went to the party, didn't they? √

2.2.6 Summary and conclusions: quantifier monotonicity and quantified sentence polarity in terms of natural language tests

While the tests discussed in this section provide a basis for the classification of quantifiers in terms of monotonicity and polarity, it is apparent that their reliance upon intuitive judgements regarding the acceptability of the resultant sentences entails that the results are not always unambiguous. However, three general conclusions would seem to arise from the application of the tests:

1. For many quantifiers, there is a consistency in the test results such that: M↓ quantifiers appear to accept NPIs and form negative sentences, while M↑ quantifiers do not accept NPIs and form positive sentences.
2. Some quantifiers (e.g. *not quite X (x)*) do not appear to be M↓, but do appear to accept NPIs and form negative sentences.
3. Some quantifiers (e.g. *at most X (x)* and *less than X (x)*) can be classified as M↓ and may accept NPI, but form positive statements by a tag test.

The issues raised here are of considerable importance and will be returned to later in the thesis.

2.3 Psychological Studies of Focus and Negativity in Quantifiers

Undoubtedly the most extensive and systematic empirical investigation of the function of natural language quantifiers has been carried out by Moxey and Sanford and their colleagues (e.g. Moxey & Sanford, 1987, 1993a, 1993b; Moxey, Sanford, & Barton, 1990; Sanford, Moxey, & Paterson, 1996). This section relates the main findings generated by this functional approach adopted by Moxey and Sanford and considers how this evidence can be best interpreted. It begins by defining the relevant terms of **anaphora** and discourse **focus**.

2.3.1 Anaphora and discourse focus

A major interest within the domain of psycholinguistic research is the study of what has been termed discourse coherence. This refers to the processes by which the relationships between consecutive pieces of text are established to create a coherent model of the discourse. Of central importance to this process is the development of isomorphic mappings between currently mentioned text entities and their antecedents in the discourse. In even the simple case (33) below, for the text to make sense *she* must be interpreted as being co-referential with *Julia* and *him* with *Bob*.

(33) Julia has decided to marry Bob because she loves him.

Here, these pronouns act as **anaphoric** references to specific characters mentioned earlier in the text. In this case the referents are unambiguously specified because the pronouns are gender marked and there exists only one possible referent of each gender. In many cases, however, the situation is more complex and there exist more than one possible referent for a pronoun. In such cases, the ease of pronominal binding to a textual antecedent has been found to be dependent on the relative prominence of the antecedent in the discourse model, that is, upon the extent to which the entity can be said to be in **focus** (e.g. Chafe, 1972; Grosz, 1977; Sanford & Garrod, 1981; Sanford, Moar, & Garrod,

1988). While, for present purposes, it is unnecessary to review the various determinants of focus, it is important to note that ease of pronominal reference can be used as an operational measure of what discourse entity is in focus at the time of the pronominal probe.

2.3.2 Anaphoric reference to the quantified plural noun-phrase

While considerable progress has been made in identifying the processes involved in resolving anaphora in cases where there exists a simple isomorphic mapping between anaphor and textual antecedent (e.g. Clifton & Ferreira, 1987; Garrod & Sanford, 1977, 1990; Gernsbacher, 1989; Stewart, Pickering & Sanford, 2000), relatively little work has focused on the more complex forms of relationship which exist. Among these is the case where the referent is a quantified noun-phrase.

The principal claim made by Moxey and Sanford (1987; 1993a) is that quantifiers which contain some negative property (these will subsequently be referred to as "negative" quantifiers) and those which appear positive in polarity, can be differentiated in terms of the set which they make available for subsequent pronominal anaphora. So while, for example, *not quite all* and *almost all* communicate similarly large proportions, the typical anaphoric referents for a noun-phrase quantified by these are different. That is, positive and negative quantifiers have different patterns of focus associated with them. This phenomenon is best illustrated by a specific example:

(34) Almost all of the football fans went to the match.

(34a) They watched it with enthusiasm.

(34b) They watched it on tv.

For (34), the continuation (34a) seems perfectly felicitous in the use of *they*, while the continuation given in (34b) does not.

(35) Not quite all of the football fans went to the match.

(35a) They watched it with enthusiasm.

(35b) They watched it on tv.

In (35), however, the continuation (35a) sounds a little strained while (35b) now sounds perfectly acceptable. To explain this, Moxey and Sanford (1987; 1993a) suggest that positive quantifiers (such as *almost all*) focus attention on what they term the *reference set* (refset) while negative quantifiers (such as *not quite all*) tend to focus attention on the *complement set* (compset). The refset being the subset for which the predicate is true (in this case, the subset of the football fans who **did** go to the match); the compset being the subset for which the predicate is not true (the football fans who **did not** go to the match). Thus, the referent for pronominal anaphora to a quantified plural noun-phrase will be dependent on the polarity of the quantifier used.

2.3.3 Quantifier focus effects in language production

In order to empirically evaluate this claim, Moxey and Sanford and their colleagues have conducted an extensive series of experimental studies. The initial study (Moxey & Sanford, 1987) made use of a constrained sentence completion methodology. In this study, 800 subjects were presented with a single sentence (of one of the forms indicated in Table 2.3) and asked to generate a sensible continuation following the pronoun *they*. This forced subjects to identify an antecedent for the pronoun, thereby providing an operational measure of the discourse entity in focus. The subjects were then asked to indicate the referent of the pronoun by ticking a category corresponding to one of: refset, compset, the set in general, all the set or "other". Independent judges then categorised the responses into these categories. Where there was disagreement between judges or between the judges and the subjects' own intuitions, the referent was classified into the "other" category.

Table 2.3 Materials used in the Moxey and Sanford (1987) study

Qx football fans went to the match **[connective]** they...

Qx MPs went to the meeting **[connective]** they....

Where:
Qx was one of: *Few*, *Very few*, *A few*, *Only a few* or *not many*.
[connective] was one of: *period* (i.e. "."), *and*, *but* or *because*.

For present purposes, only the data from the *period* connective conditions will be reviewed at this point. Evidence would seem to indicate that other forms of connective may have specific focus altering properties associated with them and evidence relating to these will be considered in some detail in later sections. Table 2.4 shows the proportion of compsets generated in this study (along with the proportion of those generated in a similar study, with different quantifiers, reported by Moxey and Sanford, 1993a). This shows a pattern of reference consistent with Moxey and Sanford's claims in that negative quantifiers give rise to a, statistically reliable, predominance of compset references while there are no compset references associated with the two positive quantifiers. However, the effect is not symmetrical in that, while negative quantifiers would seem to preferentially allow (or *license,* in Moxey and Sanford's terminology) compset reference, they do not prevent refset reference; whereas, positive quantifiers would appear to effectively prevent the possibility of compset reference. This evidence suggests that negatives are associated with a more diffuse pattern of focus.

Table 2.4 Compset reference pattern from Moxey and Sanford (1987, *1993a) for the *period* connective conditions.

Quantifier	Proportion Compset Reference
Few (x)	62.5%
Very few (x)	72.5
Not many (x)	79%
*Not quite all (x)	87%
*Hardly any (x)	86%
A few (x)	0%
*Many (x)	0%
Only a few (x)	5%

While this evidence would seem to suggest a focusing effect driven by quantifier polarity, the study was non-optimal in certain respects. Firstly, it employed only two basic scenarios and therefore the extent to which the effect would generalise across wider contexts remained unknown. Secondly, the investigation of negative quantifiers was confined to ones for which the compset was large relative to the refset (i.e. >50%). Thus compset size was not eliminated as a potential factor in the observed focus effect. Finally, the materials employed constructions which did not necessarily partition the intended superset. For example, in the first material given in Table 2.3, the intended superset is the set of football fans, however the superset could equally be interpreted as being, something like, "all people who went to the match". Here, there would seem to be a wider range of potential antecedents for pronominal reference (e.g. non-football fans who were at the match), than would be the case for materials which unambiguously partitioned a single, explicitly introduced, superset.

These issues were addressed in two sentence completion studies reported by Sanford, Moxey and Paterson (1996). These employed a wider range of materials (30 experimental items) than were used in the original Moxey and Sanford (1987) study. Additionally, the quantifiers used covered a wide range of proportions, allowing the relative size of the compset to be systematically varied across the conditions. The materials in these studies were similar in form to that of the previous study, except that each introduced an explicitly partitioned superset, i.e. each corresponded to the form *Qx of the (x)...* (see Table 2.5 for examples). A total of 600 subjects each completed a single sentence beginning with the plural pronominal probe *They*. As before, subjects then indicated the referent of the pronoun.

Table 2.5 Examples of the materials used by Sanford, Moxey and Paterson (1996)

Qx of the accountants took up the retraining offer. They....
Qx of the students produced good essays. They....
Qx of the holiday makers lay on the beach. They....

The results of this study (see Table 2.6) supported previous findings by again indicating a predominance of compset references in sentence completions for negative polarity quantifiers and an almost complete lack of these for positive quantifiers.

Table 2.6 Reference patterns (in %) from Sanford, Moxey and Paterson (1996)

	Reference Class		Polarity
Quantifier	Refset	Compset	
Not quite all	13.3	63.3	
Not all	6.6	80	
Less than half	23.3	60	
Not many	6.6	83.3	
Few	26.6	66.6%	
Not quite 10%	30	50	Negative
Not quite 30%	26.6	60	
Not quite 50%	6.6	76.6	
Not quite 70%	40	43.3	
Not quite 90%	40	33.3	
Nearly all	96.6	0	
Almost all	90	0	
More than half	83.3	0	
Many	100	0	
A few	96.6	0	
Nearly 10%	83.3	3.3	Positive
Nearly 30%	90	0	
Nearly 50%	90	0	
Nearly 70%	83.3	0	
Nearly 90%	86.6	0	

The robustness of the phenomenon is demonstrated by the persistence of the observed reference pattern across compsets of different relative sizes and in the context of more varied stimulus materials.

2.3.4 Quantifier focus effects in language comprehension

A similar pattern of differentiated focus between positive and negative polarity quantifiers has been demonstrated in the context of discourse comprehension. Sanford, Moxey and Paterson (1996) conducted a self-paced reading study using a series of experimental items (see Table 2.7 for an example) devised to manipulate the focal pattern observed in the previous completion studies. The rationale behind this study being that anaphoric references to the focused plural set would be more easily (and, therefore, more rapidly) integrated into the discourse model than anaphors to the unfocussed set. So, in light of the evidence from the completion studies which suggested that negative quantifiers favour compset reference and positives refset reference, it was predicted that reference to a compset should be more rapidly processed (i.e. integrated into the discourse model) when the referent is a negative, as opposed to a positive, quantifier (the reverse pattern holding for refset anaphors). Furthermore, because negatives appear to have a more diffuse pattern of focus associated with them, it was predicted that refset reference in the context of negative quantification would be faster than compset reference in the context of positive quantification.

Table 2.7 Example material from the Sanford, Moxey and Paterson (1996) self-paced reading study.

Writing an essay
The literature students were set an essay on Samuel Beckett and modern literary theory. **Qx** of the students produced readable results. Their **[clarity/confusion]** demonstrated how well they understood the topic.
Where *Qx* was one of: *A few, Few, Many* or *Not many*

The experimental materials for this reading time study were devised so as to introduce a (positive or negative) quantifier which partitioned a plural set, this set subsequently being the subject of an anaphoric reference contained in the following sentence. The sentence which contained the anaphor was constructed such that it described a property which was congruent with either the refset, that is, the set for which the predicate was true (*clarity* in the above example) or the compset, that is, the set for which the predicate was false

(*confusion* in the above example). Reading times for the sentence containing the anaphor were then analysed as an index of the set in focus.

Figure 2.1 Reading times (in seconds), with standard error bars, for the target sentences from Sanford et al (1996) (Reproduced from Sanford et al, 1996).

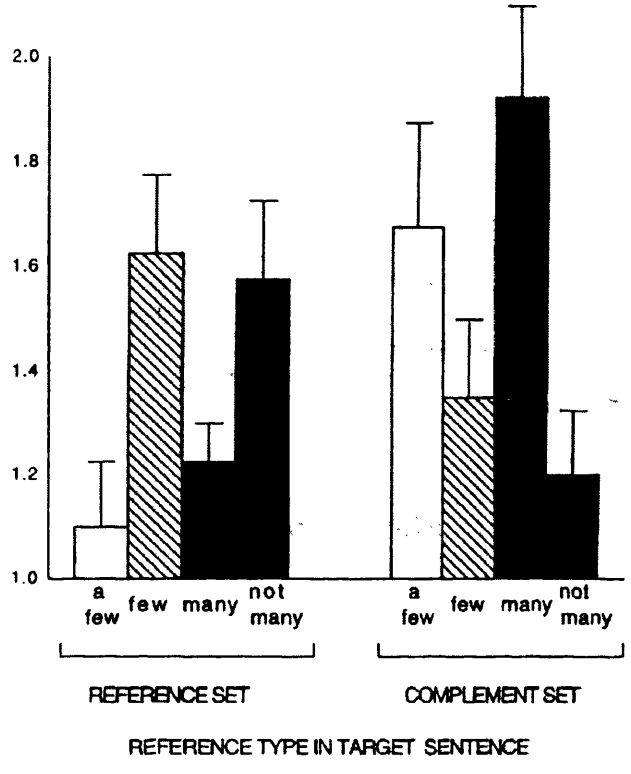


Figure 2.1 shows the reading times for the anaphoric reference sentences as a function of quantifier and reference type. The results confirmed the main hypothesis by indicating reliably faster reading times for the anaphoric reference sentences containing properties which were congruent with the predicted set focus. That is: refset congruent sentences were processed faster when the referent was a positive quantifier (*A few x* or *Many x*) than when the referent was a negative quantifier (*Few x* or *Not many x*) and compset congruent sentences were read faster when the referent was a negative quantifier. This evidence supported previous findings in again demonstrating the association of positive quantifiers with refset focus and negative quantifiers with compset focus.

2.4 Unresolved Issues: Licensing Conditions and Variable Rates of Compset Reference

While the empirical evidence related above would seem to demonstrate the association of diffuse or compset focus with some negative property of the quantifier involved, it is evident that a full description of the effect requires further specification. In respect to this, two major issues remain unresolved. First, what precise property of negativity is it which gives rise to the effect? Secondly, there exists a considerable degree of variance in the rate of compset reference associated with different quantifiers. This would seem to suggest that whatever property of the quantifier it is which is responsible for the focusing effect, this property is not uniformly present across all quantifiers which give rise to the effect. This section briefly considers these issues. Table 2.8 provides a summary of pertinent information concerning judgements of quantifier monotonicity and polarity and variable rates of compset reference.

2.4.1 Licensing conditions for compset focus

In considering the conditions which license complement set reference, two principal candidates are evident: quantifier downwards monotonicity and quantified sentence negativity. As discussed earlier in this chapter, intuitive judgements based on natural language tests of monotonicity and sentence polarity generally yield consistent results (M↓ quantifiers appear to form sentences which accept NPIs and are judged to be of negative polarity by tag tests).

Kibble (1997) has suggested that the property of quantifiers which licenses complement set anaphora may be downwards monotonicity. While Kibble's position will be evaluated in full in Chapter 3, it is pertinent to examine his main contention at this point. On the surface, this would seem to be a relatively easy claim to evaluate. If the quantifiers associated with compset reference appear to be M↓ in terms of the tests previously discussed, the claim is supported. If they do not the claim is falsified. However, a major

problem in the resolution of this issue is that tests of monotonicity do not always give consistent answers. Recall that Zwarts (1996) has suggested that M↓ quantifiers combine grammatically with negative polarity items. Thus, in principle, quantifiers which accept negative polarity items should also be M↓ in sentences which test basic monotonicity and, if Kibble is correct, should license complement set anaphora. This, however, does not always seem to be the case. For example *not quite all* is associated with a high rate of compset reference (see Table 2.6) and while it does not seem to be M↓ by sentence (36), it does appear to accept negative polarity items as in sentences (37) and (38):

(36) If *not quite all* of the fans went to the concert, then *not quite all* of the fans went to the concert early.

appears to be a false assertion, while

(37) *Not quite all* of the students take drugs anymore.

and

(38) *Not quite all* of the students give a damn about lectures.

seem to be reasonably acceptable.

2.4.2 Variability in the rate of complement set reference

A second issue which remains to be resolved concerns the variable rates of compset reference obtained in the context of different quantifying expressions (as illustrated by Table 2.8). While some form of negativity (as indexed in terms of either the acceptance of NPIs or positive tag questions) would seem to be associated with the occurrence of complement set reference, no theoretical mechanism has been suggested which would account for the range of compset reference observed in the completions.

Table 2.8 Quantifier monotonicity and polarity in terms of natural language tests and the associated rate of compset reference from Sanford et al (1996)

Quantifier	Monotonicity	Negative Polarity Item	Tag Test	% Compset Reference
Not quite all	?	yes	-ve	63.3
Not all	M↓	yes	-ve	80
Less than half	M↓	yes?	+ve	60
Not many	M↓	yes	-ve	83.3
Few	M↓	yes	-ve	66.6
Not quite 10%	?	yes	-ve	50
Not quite 30%	?	yes	-ve	60
Not quite 50%	?	yes	-ve	76.6
Not quite 70%	?	yes	-ve	43.3
Not quite 90%	?	yes	-ve	33.3
Nearly all	?	no	+ve	0
Almost all	?	no	+ve	0
More than half	M↑	no	+ve	0
Many	M↑	no	+ve	0
A few	M↑	no	+ve	0
Nearly X%	?	no	+ve	0-3.3

Where:
Monotonicity refers to the acceptability, or otherwise, of the quantifiers substitution into the following sentences:
If *Qx* of the fans went to the concert early, then *Qx* of the fans went to the concert. (M↑).
If *Qx* of the fans went to the concert, then *Qx* of the fans went to the concert early. (M↓).
M↑ indicates monotone increasing; M↓ indicates monotone decreasing and "?" indicates nonmonotone or undecidable results.
Negative Polarity Item refers to the grammatical combination of a quantified statement with a negative polarity item. "?" indicates mixed or undecidable results.
Tag Test refers to a quantifier polarity judgement in terms of the acceptability of the quantified statements' combination with positive and negative tag questions. "?" indicates undecidable results.

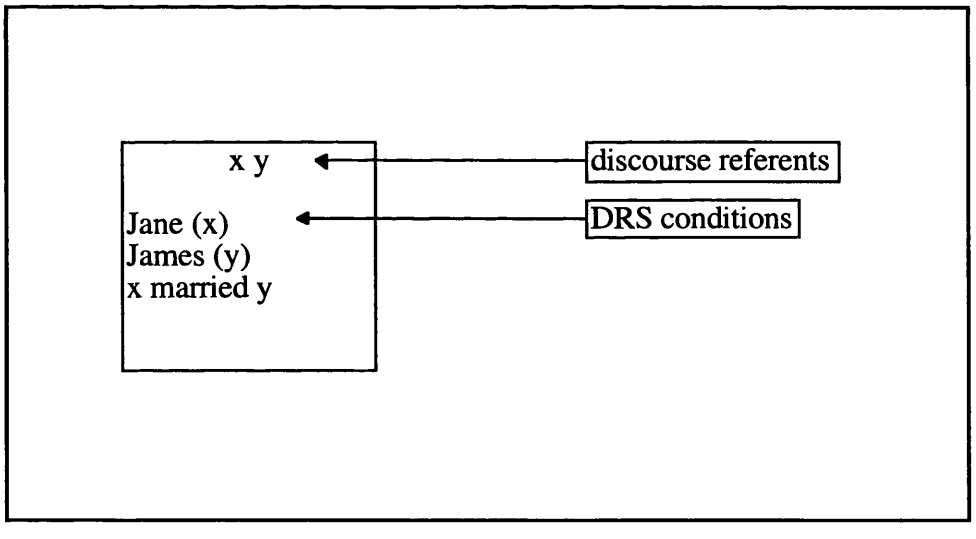
This observation is important because it would seem to suggest that there may exist some other property of the quantifying expression which predicts the strength of compset focus. Given that some aspect of negativity would seem to be the best predictor of compset focus, it is possible that an additional property associated with this may explain the variable rates. This possibility is explored in more detail in Chapter 3.

2.5 Discourse Representation Theory and Quantifier Focus

Kamp and Reyle's (1993) Discourse Representation Theory (DRT) provides a semantic framework for the representation and interpretation of discourse. Within this theoretical framework, Discourse Representation Structures (DRSs) are formed and dynamically updated during discourse processing. The DRS formed for a given piece of discourse is composed of two elements: the available referents of the discourse and the specified DRS conditions. Conventionally, these are related in the form of a diagram (such as that given in Figure 2.2), with the referents being given at the top and the conditions immediately beneath these. Given a sentence such as (39) below, a DRS such as that given in Figure 2.2 is held to result. Here both "Jane" and "James" are available for subsequent pronominal reference.

(39) Jane married James.

Figure 2.2 DRS for "Jane married James"



DRT also allows for subsequent plural pronominal anaphora where the antecedent is a previously introduced set, or where the referent can be computed (either by *summation* or *abstraction*) from previously introduced text entities. The process of summation involves the combination of previously mentioned referents (in the above example, these would correspond to Jane and James) into a single set. This new referent (which represents the

union, in Kamp and Reyle's terminology, of the previously available referents) can now be referred to by a plural pronoun. The process of abstraction provides a mechanism which, given a sentence containing a plural quantified noun-phrase, allows for certain sets/subsets to be available for subsequent pronominal anaphora. For example, in sentence (40) below,

(40) Many of the students at Glasgow University enjoy writing essays.

the three sets available for pronominal reference would correspond to:

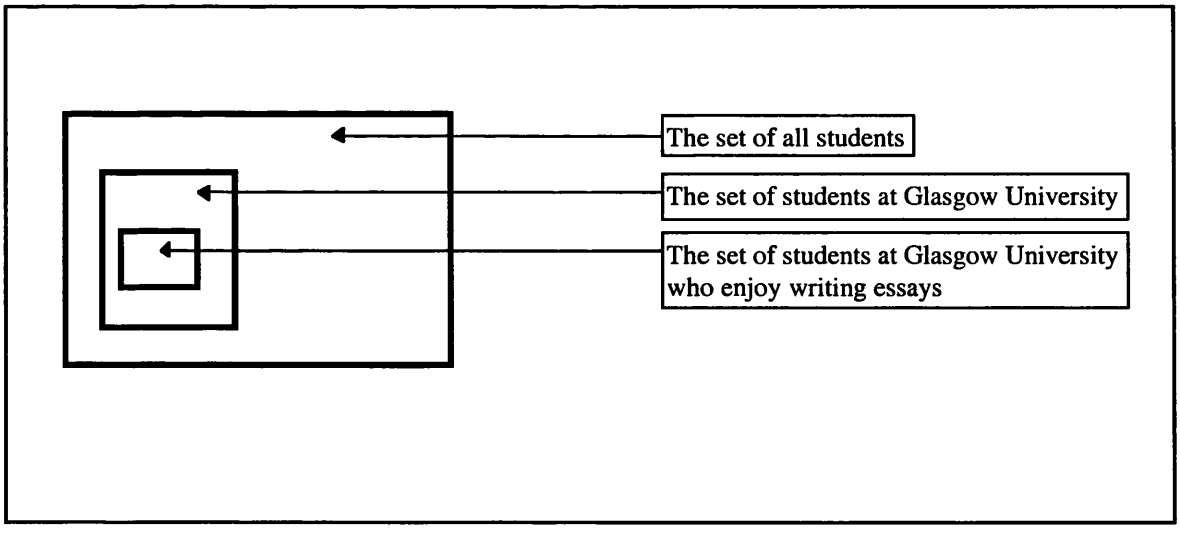
1 - the set of all students.

2 - the set of students at Glasgow University.

3 - the set of students at Glasgow University who enjoy writing essays.

These sets are illustrated by Figure 2.3 below.

Figure 2.3 The three sets available for pronominal reference for the sentence: "Many of the students at Glasgow University enjoy writing essays" under DRT.



Similarly, the sets available for anaphora in quantified sentences of the form used by Moxey and Sanford, such as (41) below, would correspond to:

(41) Many of the football fans came to the match.

1 - the set of all football fans.

2 - the set of football fans who came to the match.

Therefore, it is evident that in the DRT framework, there exists no set abstraction which would correspond to a compset. This has led Corblin (1996) to suggest that compset reference is only apparent, rather than actual. Corblin uses the "marbles" example (42) as a linguistic illustration of the impossibility of compset reference:

(42) Eight of ten marbles are in the bag. They are under the sofa.

In this example, it is only possible to generate the compset by way of subtracting the refset (the eight marbles) from the superset (the ten marbles). The apparent impossibility of interpreting the plural pronoun as an anaphor to the compset is taken to indicate that compset reference is not possible. However, as Moxey and Sanford (2000a) remark, because *eight of ten marbles* is a positive quantifier, it would not be predicted that the compset would, in any case, be in linguistic focus. If, they argue, a negative quantifier were used (e.g. *few of the ten marbles*), pronominal reference to the compset would seem possible. However, it should be noted that even in this case, compset reference sounds a little strained.

Arguing from the standard DRT position, Corblin (1996) has suggested an alternative interpretation of Moxey and Sanford's findings. Corblin's main claim is that cases of apparent compset reference can best be explained in terms of reference to some generalised set (termed the *maxset*) which is available within the DRT framework. This position entails that the pronoun is interpreted, rather than as referring to the compset, as referring to some aspect of the superset in general. So, in (43), *They* refers to the set of students in general.

(43) Not many of the students went to the pub. They *(generally)* got drunk at home instead.

Moxey and Sanford (2000a), while accepting that generalisations of this sort will occur in some instances, provide a series of specific arguments which suggest that compset reference cannot always be explained in this way. First, they argue that, in the completion studies, participants were required to specify the referent of the pronoun (from, as already described, categories which included both the compset and the set in general). The fact that participants' responses and raters' intuitions coincided in the majority of instances is taken as indicative of the actuality of compset reference. Moreover, Moxey and Sanford indicate that there is no evidence to suggest that the participants misinterpreted the nature of the task or had any difficulty specifying the referent. The second argument they put forward relates to the absence of any explicit terms in the completions which would signify generalisations. Terms such as *typically, mostly* or *generally* are, they observe, notably absent from the completions. Thirdly, they note the presence of expressions in the completions which unambiguously applied to members of the compset (and could not be taken to refer to members of the refset). An example continuation of this type, termed here an *exclusive-set* completion, is given in (44).

(44) Not quite all of the football fans went to the game. They couldn't get to the park because of the train strike. (Moxey and Sanford, 2000a, p331).

Here, the completion can only apply to the fans who were not at the match (i.e. the compset). Finally, they note that compset reference is evident even in cases where the compset corresponds to a very small proportion of the superset (as in *Not quite all*), which, it is suggested, makes generalisation less plausible. These arguments would seem to suggest that a process of generalisation is insufficient explanation for all occurrences of apparent compset focus.

2.6 Chapter Summary and Conclusions

This chapter began by defining and attempting to classify quantifiers in terms of monotonicity and polarity. A substantial body of empirical evidence was then reviewed which suggested that quantifiers could be characterised according to linguistic set-focusing properties associated with these. Specifically, some aspect of downwards monotonicity or negativity would seem to license diffuse or compset focus, while upwards monotonicity/positivity is associated with refset focus. However, as a result of the intuitive nature of judgements concerning the monotonicity and polarity of quantifiers, the precise nature of the licensing conditions for diffuse/compset focus remains unspecified. Finally, the chapter reviewed an alternative interpretation of the empirical data and concluded that this was insufficient to account for the observed phenomenon of compset focus.

Chapter Three

The Theoretical Basis of Complement Set Reference

3.1 Introduction

In Chapter 2, evidence was presented which clearly indicated the existence of a set-focusing effect associated with quantifier polarity (refset reference being linked with positively quantified statements and compset reference with negative statements). A major issue which remains unresolved, however, concerns the mechanism through which these focus effects arise. For, while the phenomenon of refset reference is currently explicable in terms of the semantic framework provided by DRT, the mechanism by which compsets may become available for pronominal anaphora is not immediately apparent. The present chapter addresses this issue and considers the mechanisms which have been advanced to explain the existence of complement set focus.

In broad terms, two specific types of model have been advanced to account for the phenomenon. The first (termed here the *Set-Driven* approach) proposes that complement sets are made directly available for subsequent pronominal reference through some type of direct mapping between the quantifier and compset. This category of approach is principally associated with attempts to explain the effect in terms of some particular semantic property of the quantifier (negativity or downwards monotonicity) which would allow compset reference. The second approach (termed the *Inference-Driven* model) suggests that a more elaborate mechanism (based on the inferences arising from the quantified sentence as a whole) is responsible for the effect.

This chapter relates the models outlined above and considers the extent to which each is consistent with the available empirical evidence.

3.2 Set-Driven Accounts of Complement Set Focus

As described in Chapter 2, the current formulation of DRT does not suggest a mechanism by which compset focus might occur. However, in the light of the bulk of empirical evidence which suggests the reality of compset focus, some theorists (e.g. Devlin, 1997; Kibble, 1997) have attempted to develop models which would allow for complement set reference within the context of quantifier or discourse semantics. This section provides an overview of the models.

3.2.1 Semantic models of complement set reference

The approach adopted by Kibble (1997) takes as its basis the alternative conceptualisation of quantifier downwards monotonicity proposed by Van den Berg (1996). Berg suggests that quantifier downwards monotonicity, rather than being directly defined in terms of itself, can be modelled in terms of negations of monotone increasing quantifiers. Kibble (1997) argues that monotone decreasing quantifiers can be interpreted as the result of one of two possible types of negation, termed external and internal negation. In the case of external negation, a given M↓ quantifier is held to be the opposite of some M↑ quantifier (designated ~q). For example, ~q for the quantifiers: *not many, less than 30%* and *not all*, would be *many, more than 30%* and *all*, respectively. In terms of sentence (1), external negation would result in the sentence being represented as something like (1b).

For internal negation, the transformation corresponds to that part of the superset not denoted by the stated quantifier (the transformed quantifier being designated q~). Thus, q~ for *less than 30%* and *not all*, would become *70% or more* and *some* respectively. Again, in terms of sentence (1), internal negation would result in the sentence being represented as (1c). Kibble notes that this type of transformation can only be performed on proportional quantifiers (i.e. those which specify a proportion of some superset, e.g.: *less than 20%* or *few*). Non-proportional quantifiers (e.g. *less than 20*) cannot be

internally negated because that part of the superset not denoted by the quantifier is impossible to determine.

(1) Less than 30% of the soldiers were overweight.

(1b) It is not the case that the set of soldiers who were overweight was a set of more than 30% of the soldiers. (Here ~q *more than 30%* is the result of the external negation of *less than 30%*).

(1c) The set of soldiers who were not overweight was a set of 70% or more of the soldiers. (Here q~ *70% or more* is the result of the internal negation of *less than 30%*).

Kibble suggests that while both types of negation result in the same truth conditions being met, the specific type of transformation employed for a given sentence will dictate which potential referents are available for subsequent anaphora. In the case of external negation (e.g. sentence (1b)), the sentence is encoded in terms of the set for which the predicate is true (i.e. the soldiers who were overweight) and the available referents would correspond to either refset or the set in general. For internal negation (e.g. (1c)), the sentence is encoded in terms of the set for which the predicate is false (i.e. the soldiers who were not overweight) and the available referents would be compset or the set in general. Thus, downwards monotonicity and the way in which this is represented become central to the phenomenon of complement set focus. Kibble, then, bases his approach upon the semantic properties of monotone decreasing quantifiers and provides a semantic mechanism which would allow for complement set anaphora.

The approach adopted by Devlin (1997) involves an expansion of the existing DRT framework. Devlin proposes the introduction of an additional rule into DRT in order to explain the phenomenon of compset reference. This rule, which she terms *Distraction*, is analogous to the existing DRT rule of Abstraction. Recall that Abstraction (as briefly

explained in Chapter 2) provides a mechanism for the generation of subset referents for a quantified sentence in terms of the predicate being true. So, given a sentence such as *Many of the football fans went to the game*, it is possible to refer pronominally to those who did go to the match. In contrast, Devlin's suggested rule of Distraction operates when the quantifier can be interpreted as negating the predicate and provides a mechanism for the generation of referents in terms of this negated predicate. For example, given *Not many of the football fans went to the game*, it is possible to pronominally refer to those who did not go to the match.

3.2.2 Problems with the semantic approaches

Both Kibble and Devlin present plausible, if apparently somewhat ad hoc, semantic solutions to account for complement sets. However, it is evident that certain issues remain problematic. In essence, these correspond to the general problems discussed in Chapter 2 (section 2.4). These are: first, the specific conditions which license complement set reference and, second, the variable rates of complement set focus obtained in the empirical work of Moxey and Sanford.

Kibble suggests that downwards monotonicity may be the property which licenses compset focus. However, as previously discussed, natural language tests for monotonicity do not appear to always yield consistent and unambiguous results. Additionally, since complement set reference does not appear to be uniquely predicted by M↓ quantifiers (e.g. the quantifier *not quite X* does not seem to pass standard set-inclusion tests for downward monotonicity, but is associated with compset reference, as is the patently nonmonotone *only a few*), Kibble's perspective does not seem to be supported by the available empirical evidence.

Devlin's contention that some form of quantifier negativity may drive compset focus is, perhaps, less obviously problematic. For, it is apparent that some aspect of negativity (as judged by intuitive judgements concerning acceptance of NPIs and/or positive tag

questions) would appear to be associated with compset reference (see Table 2.8 for an illustration of the general relationship between monotonicity, negativity and the occurrence of compset focus). However, as she admits, the property of negativity itself is not fully understood and it remains unclear in her description just why this property should trigger Distraction (and therefore allow compset reference) on some occasions but not in others. This failure to approach an explanation of the variability in complement set rates obtained in empirical studies is common to both of these semantic models and represents a major problem for the semantic accounts as they are currently formulated.

3.3 Negativity and the Development of the Inference-Driven Model

In common with the account proposed by Devlin (1997), the initial explanatory mechanism for compset focus suggested by Sanford, Moxey and Paterson (1996) focuses upon the role of negativity in the generation of complement set reference. Unlike Devlin's semantic account, however, the conceptual basis for the Inference Model has developed from previous work concerning the processing, representation and function of negative statements. This section reviews the pertinent findings of this previous research in order to provide a conceptual context for the Inference Model.

3.3.1 The processing of negative statements

A substantial body of research (e.g. Just & Carpenter, 1971; Trabasso, Rollins & Shaughnessy, 1971; Wason & Jones, 1963) has indicated that negative statements are relatively harder to process (as indexed by the response latency in sentence verification tasks) than corresponding positive sentences. The basic task of the participant in such a study is to provide a judgement of the truth or falsity of a target sentence in relation to either world knowledge or a second stimulus. The example target sentences in (2) depend on world knowledge and relate statements concerning the property of *oddness* or *evenness* of a specified number. The target sentence can, as related by Horn (1989),

appear in one of 4 possible permutations: *Affirmative sentence/True* (a), *Affirmative sentence /False* (b), *Negative sentence/True* (c) or *Negative sentence/False* (d).

(2) (a) 7 is an odd number.

(b) 6 is an odd number.

(c) 6 is not an odd number.

(d) 7 is not an odd number.

Two general conclusions are evident from such research. First, as indicated above, the time taken by participants to assess the truth or falsity of a statement is shorter for affirmative statements, such as (a) and (b), than for negative counterparts, such as (c) and (d). A second conclusion evident from such studies (e.g. Clark, 1976; Wason, 1961) is that there exists a progressive increase in average response latency through the various target sentence configurations, such that: the affirmative-true condition yields the fastest response, followed by affirmative-false, followed by negative-false, with negative-true sentences being responded to slowest of all. This finding indicates that while *true* responses are faster for positive statements, the reverse is the case for negative statements (*false* responses being faster than *true*).

Explanations for these findings have been offered by Clark (e.g. 1976) and Wason (e.g. 1971) respectively and these are reviewed in the forthcoming sections.

3.3.2 The encoding of negative statements

Horn (1989), in reference to empirical work by Fodor and Garrett (1966) and Just and Carpenter (1971), suggests that the finding that negative statements are associated with longer verification times is not sufficiently explained in terms of either the increased processing load resulting from the relative complexity of the rules which govern negative placement, nor from the fact that negative sentences are generally longer than positives. A more plausible, and empirically supported, explanation for this phenomenon has been

offered by the "True" model proposed by Clark (1976). The "True" model attempts to explain the increased processing time associated with negative statements in terms of the way in which negative statements are mentally represented. Negative propositions, Clark suggests, can be represented in terms of their positive counterparts. This can be illustrated in reference to the sentence given in (3).

(3) Jane did not go to the gallery.

Here, the "True" model suggests that this proposition might be represented as something like *false(Jane went to the gallery)*. Consequently, since the transformation of a negative proposition into its affirmative counterpart necessitates an additional processing step over and above that required for an affirmative proposition, negative propositions will take longer to process (assuming that the processing is serially enacted).

While Clark's position suggests that there exists a specific difference in the way simple negative and affirmative statements are represented, it is of some considerable importance to the subject of this thesis to ascertain if a similar argument might hold in the context of positively and negatively quantified statements. Just and Carpenter (1971) investigated the way in which quantified statements are processed in the context of a picture-sentence verification task. Participants were presented with a square matrix composed of 16 coloured dots (each of which could be either red or black) and asked to assess the truth or falsity of a quantified sentence pertaining to the number of dots of a given colour. The experimental manipulation of interest here involved the comparison of verification times for positively and negatively quantified statements. Here, the participant would be shown a matrix of dots and asked to assess the truth or falsity of a sentence of the form *Qx of the dots are red* or *Qx of the dots are black* (where *Qx* could be positive or negative). So, for example, given a matrix composed of 3 red dots and 13 black dots, both the statements *Few of the dots are red* and *Many of the dots are black* would match with the picture (and yield a *true* response), while reversing the quantity statements in the sentences would create a mismatch (*false* response).

Just and Carpenter found that verification times for the positive condition were shorter than those for the negative condition. Additionally, they found evidence to suggest that there existed differences in the way the positively and negatively quantified statements were encoded in the context of this task. For the positives, the sentence was interpreted in terms of the colour of the dots explicitly mentioned in the sentence (which would correspond to the refset, in the terminology adopted in this thesis). So, for example, in the sentence *Many of the dots are red*, the problem would be encoded in terms of the red dots. However, for the negative sentences, it was found that the sentence was encoded in terms of the dot colour not explicitly mentioned (corresponding to the compset). So *Few of the dots are red* would be encoded in terms of the black dots (which constitute the smaller subset).

This finding is important in two respects. First, it provides further support for the reality of the differential focus effect associated with positive and negative quantifiers from the standpoint of a different methodological approach. Secondly, any theoretical mechanism advanced to explain the focus effect evident in these data is likely to be similarly relevant to the Moxey and Sanford data. Clark (1976) has argued that the results of this study are explicable in terms of the "True" model, the argument being that: given a sentence such as *Not many of the dots are red*, this is encoded as *(false (many (dots are red)))*, the verification procedure then involves testing what it is that is true of *many*, which consequently focuses the participant on the black dots (Sanford et al, 1996).

3.3.3 The function of negative statements

An additional strand of empirical evidence concerning the processing of negativity suggests that the time taken to process negative statements varies as a function of the context in which they occur. This idea has been most notably expressed by Wason (1965) in terms of what he calls the *contexts of plausible denial* . Wason's principal contention is that a primary function of negativity may be to signal that the information communicated violates some existing expectation or presupposition. Thus, statements which deny a state

of affairs are more acceptable (as indexed by verification latencies) if they occur in a context within which there is reason to believe that the contrary may have been expected by someone. Conceptually, this can be illustrated in terms of sentences (4) and (5).

(4) That tree doesn't lose its leaves in the winter.

(5) That tree doesn't lose its roots in the winter.

Here, sentence (4) may be intuitively more acceptable than sentence (5) because it is plausible that someone may not have realised that the tree in question was not deciduous. However, it is unlikely that anyone would presuppose that any tree might shed its root system during winter. Likewise, a statement such as *6 is not an odd number* may be relatively difficult to process because there is little scope for interpreting this as violating some existing expectation. Thus denial would seem to be appropriate only when there are grounds to believe that some contrary state of affairs may have been expected.

3.4 The Inference-Driven Model

An alternative explanatory mechanism for complement set reference to that of the semantic level accounts proposed by Kibble and Devlin, has been suggested by Moxey and Sanford (e.g. Moxey & Sanford 2000a; Sanford et al, 1996). This account, termed here the *Inference-Driven Model*, proposes that the phenomenon of compset focus is principally driven by the kinds of inferences which typically arise in response to a negatively quantified statement. In actual fact, the Inference Model has been subject to some degree of refinement and development since it was initially advanced. This section briefly introduces the two versions of the model.

3.4.1 The Inference Model: Version one

According to the initial form of the model proposed by Sanford et al (1996), quantifiers which form negative statements (when combined with a simple declarative sentence) will

tend to make the negation of the sentence predicate relatively prominent within the discourse model. So, for sentence (6), emphasis will tend to fall on *not* going to the match. If this is the case, the inferences subsequently generated by the reader may principally concern the negated predicate (specifically they might concern the reasons for, or consequences of, some state of affairs not being the case). Under such circumstances, compset reference may result because the inferences generated concern the set which is associated with the predicate negation (i.e. the compset corresponds to the set of fans who did not go to the match).

(6) Not many of the fans went to the match.

3.4.2 The Inference Model: Version two

The current version of the model (Moxey and Sanford, 2000a) further specifies the role of negation in driving inferential activity and, consequently, compset reference. In this form of the model, it is the aspect of presupposition denial implicit in negative statements which guides the pattern of inferences. Following on from the work of Clark (1976) and Wason (1965), a negatively quantified statement can be held to introduce a presupposition and then deny it. Thus, in terms of the negative statement in (6) above, the presupposition that *many* may have been believed to be the case is introduced and then denied. Negatively quantified statements can therefore be argued to signal some form of expectation violation. Subsequent inferential activity may then be anchored to this anomalous occurrence in an attempt to explain, justify or otherwise comment upon the unexpected state of affairs. Again, since the set associated with the expectation violation corresponds to the compset, it would be expected that the inferences generated would typically refer to, or somehow involve, this set. In terms of the quantified sentence completion and comprehension studies previously reviewed, the pronominal anaphor encountered subsequent to the expectation violation signalled by the negatively quantified sentence, would then be interpreted as referring to the compset.

In summary, the Inference-Driven Model suggests that negatively quantified statements may signal an expectation violation and subsequently trigger a type of inferential activity which attempts to explain or justify this unexpected state of affairs. Since these inferences may typically concern the compset, complement set anaphora is licensed.

3.5 Sentential and Constituent Negation

A central aspect of the inference model is that negatively quantified statements can be held to assert a denial of some existing presupposition. Thus, it is important to note at this point that not all negative statements are considered to signal denial. According to Clark (1976), negative statements can be separated into two categories: those which express *Sentential negation* (S-negation) and those which express *Constituent negation* (C-negation). Clark claims that it is only S-negation which is associated with denial (C-negation being linked with an affirmation of a negative state of affairs). This section relates a series of diagnostic tests for these properties and attempts to use the concepts to resolve a classification problem identified in Chapter 2. Finally, the section considers how the results of these diagnostic tests may provide the potential for an explanation for the variable rates of compset focus which have been observed to be associated with different quantifiers.

3.5.1 Klima's diagnostic tag sentences reinterpreted

Chapter 2 introduced a natural language test of negativity which was based upon Klima's (1964) observation that simple negative statements could be grammatically combined with positive confirmatory tag questions (e.g. *did they?*) and positive statements with negative tags (e.g. *didn't they?*). Horn (1989), however, claims that tag tests such as these do not simply serve to separate positive and negative linguistic environments, rather they can provide a diagnostic tool for the separation of affirmations from denials. Thus, the properties of S- and C-negation can be indexed with reference to these tests. In addition to

positive versus negative confirmatory tag questions, Horn relates a further three types of diagnostic tag sentences which derive from Klima's work. These are introduced below.

The first of these (7) (Horn, 1989 p184) involves the use of *either* versus *too* tags. Here, *either* appears to acceptably connect two denials and thus indicates S-negation. The use of *too* is, however, acceptable for affirmative statements (indicating C-negation).

(7) Mary isn't happy and John isn't happy *either.* √

Mary isn't happy and John isn't happy *too.* ×

Mary is unhappy and John is unhappy *either.* ×

Mary is unhappy and John is unhappy *too.* √

A similar pattern of results is indicated in terms of *neither* versus *so* tags. In (8) (Horn, 1989 p184) *neither* is acceptable for S-negation and *so* for C-negation.

(8) Mary isn't happy and *neither* is John. √

Mary isn't happy and *so* is John. ×

Mary is unhappy and *neither* is John. ×

Mary is unhappy and *so* is John. √

The final diagnostic test involves the acceptability of a *not even* tag. The sentences in (9) (Horn, 1989 p185) indicate that *not even* is acceptable in the context of S-negation, but not C-negation.

(9) The attacks weren't successful, not even the last one. √

The attacks were unsuccessful, not even the last one. ×

3.5.2 An explanation for the observed dissociation between downwards monotonicity and negativity found in some quantifiers

This finer grain classification of negative statements in terms of their either representing a denial (S-negation) or an affirmation (C-negation) would seem to potentially provide the conceptual basis for the resolution of a problem identified in Chapter 2. Recall that while many M↓ quantifiers (e.g. *not many (x)*) appeared to be negative in polarity (as indexed by both the acceptance of NPIs and positive confirmatory tag questions), others (e.g. *at most x (x)*) appeared to be negative as judged by the acceptance of NPIs, but appeared to form positive statements in that they accepted negative tag questions. In terms of the classification scheme outlined above, this apparent inconsistency between the results from different indices of negativity would appear to be due to the tests being sensitive to different aspects of negativity. Thus, *not many (x)* accepts NPIs (showing that it is M↓ and forms negative statements) and positive tag questions (showing that the quantifier is associated with S-negation and therefore forms a statement of denial). However, *at most x (x)*, accepts NPIs but not positive tag questions and thus appears to be M↓, negative in polarity and associated with C-negation. This differential classification of *not many (x)* and *at most x (x)* is also demonstrated by the other tag tests outlined above.

In sentences (10), *not many (x)* appears to accept *either, neither* and *not even* tags. However, in sentences (11), *at most x (x)* accepts *too* and *so* tags but not the *not even* tag. Thus *not many (x)* would appear to be a negative expression which forms statements of denial, whereas *at most x (x)* would appear to be a negative expression which forms statements which express C-negation.

(10) Not many men are happy and not many women are happy *either*. √

Not many men are happy and not many women are happy *too*. ×

Not many men are happy and *neither* is Mary. √

Not many men are happy and *so* is Mary. ×

Not many of the candidates were good, *not even* John. √

(11) At most 10 men are happy and at most 10 women are happy *either.* ×

At most 10 men are happy and at most 10 women are happy *too.* √

At most 10 men are happy and *neither* is Mary. ×

At most 10 men are happy and *so* is Mary. √

At most 10 of the candidates were good, *not even* John. ×

The differential classification of negative quantifiers in respect of S- and C-negation is not, however, always consistent in terms of these tests. For example, in sentences (12) which index S-negation, *less than x (x)* appears to form S-negations by accepting *neither* and *not even* tags but not positive confirmatory or *either* tags. The pattern is similarly mixed for sentences (13) which measure C-negation.

(12) Less than 10 men went to the party, did they? ×

Less than 10 men are happy and less than 10 women are happy *either.* ×

Less than 10 men are happy and *neither* is Mary. √

Less than 10 of the candidates were good, *not even* John. √

(13) Less than 10 men went to the party, didn't they? √

Less than 10 men are happy and less than 10 women are happy *too.* √

Less than 10 men are happy and *so* is Mary. ×

Thus, *less than x (x)* appears to form denials as indexed by some tests, but not as indexed by others. This matter is further complicated by the fact that these judgements are based upon intuitive acceptability rather than objective measurement. These issues raise the possibility that it may not always be possible to unambiguously decide upon the occurrence of a denial. Consequently, it would seem reasonable to suggest that the property of denial may potentially be best described in terms of a continuum of possible strengths, rather than as an absolute category. While this intuitive speculation would appear to be plausible, the matter can only be fully resolved through an empirical investigation of the variability in judgements of denial.

3.5.3 The implications of denial detection ambiguity for the Inference Model

The Inference Model states that denial may be the property of negativity which triggers complement set focus. If it is not always possible to unambiguously detect the occurrence of a denial, then any process which depends upon the detection of denial may be sensitive to this ambiguity. If this is the case then, in terms of the Inference Model, the rate of compset reference which is associated with a specific quantifier should be sensitive to the ease with which that quantifier can be held to assert a denial. This possibility would appear to present the basis for a possible explanation of the variable rates of complement set reference which have been observed in quantified sentence completion tasks. For, if quantifiers which license compset reference can be separated with respect to the extent to which they are perceived to signal denial, this variable strength of denial may potentially predict the pattern of compset reference rates between different quantifiers. This matter is addressed, on a relatively small scale, by Experiment 1.

3.6 Empirical Evidence for the Inference Model

This section reviews several key strands of evidence which provide empirical support for the Inference-Driven Model. It begins by evaluating the claim that negatively quantified

statements signal the denial of a presupposition. It then considers evidence from the types of continuation which are typically generated in response to a quantified sentence completion task. Finally, it examines the impact of connective upon reference pattern and content types generated in such studies.

3.6.1 Quantifiers and expectation violation

As related above, a key function of negatively quantified statements may be to signal some form of expectation violation. This aspect of quantifier function was directly tested empirically by Moxey and Sanford (1993b). Two of the experiments reported are of principal importance in this context. The first investigated the impact of quantifier upon the actual quantity judged to be communicated in a given situation (Moxey and Sanford term this a "Level 1" interpretation). The second investigated the impact of quantifier upon the quantity previously expected by the communicator (this judgement of communicator expectation was termed a "Level 2" interpretation). Participants in both studies were each presented with one of three basic scenarios (each associated with a previously established baserate expectation) along with a quantified statement. The participants were then asked to provide a quantity estimate in relation to this. In the first experiment, the task of the participant was to provide a percentage estimate of the quantity communicated. In the second experiment, the task was to provide an estimate of the percentage that the writer had expected prior to knowing the actual quantity associated with the scenario.

Table 3.1 illustrates the three materials used and relates the associated baserate expectations. The principal argument advanced was that if the quantifiers did not signal expectancy information, there should be no systematic difference between the estimates generated for Level 1 and Level 2 interpretations. That is, the quantities judged to be communicated and expected should not differ for a given quantifier if there is no expectancy information implicit in the quantifier.

Table 3.1 Materials used in the Moxey and Sanford (1993b) study

"The residents association's Xmas party was held last night in the town hall. ***Qx*** of those who attended the party enjoyed what might be called the social event of the year." (66%; high base rate)

"At yesterday's party conference, Mr. Cameron spoke about the effects of education cuts on British Universities. ***Qx*** of his audience were convinced by his conclusions." (50%; medium base rate)

"Survey has recently been carried out to find out whether or not female students prefer to be examined by female doctors. ***Qx*** of the local doctors are female" (27%; low base rate)

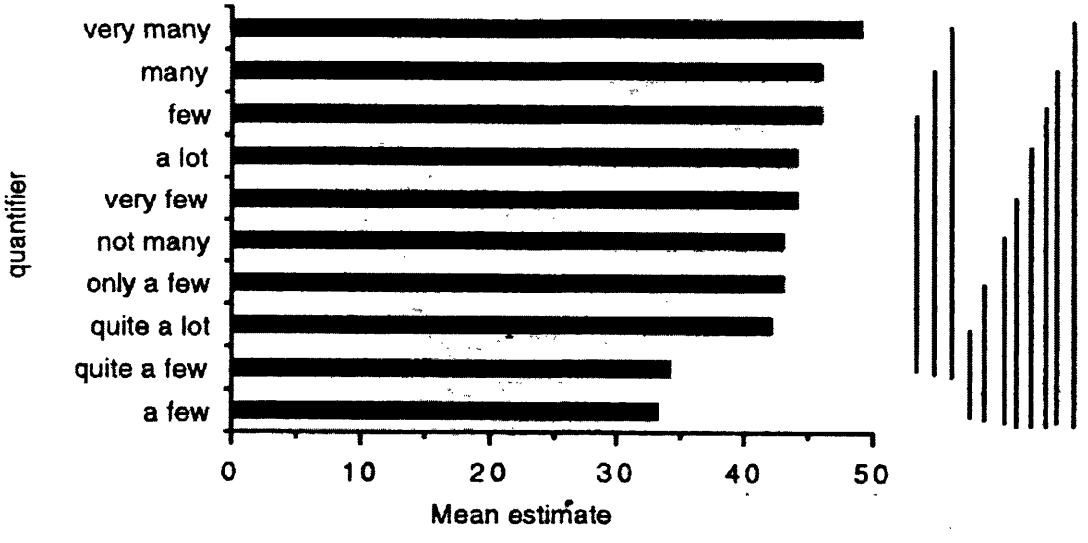
Where *Qx* was one of:
a few, quite a few, a lot, quite a lot, many, very many, few, only a few or *not many.*

Mean baserate estimates are shown in parenthesis.

For the Level 1 study, there was no reliable difference in the quantity estimates generated by participants for the low ranking quantifiers *a few, very few, few, not many* and *only a few.* For level 2, however, the pattern was rather different. Figure 3.1 shows the mean estimates, collapsed across scenarios, obtained in the Level 2 experiment.

The principal finding of interest here concerned the quantifier *a few* (which, intuitively, appears to form affirmative statements in terms of natural language tests). This quantifier signals a relatively low amount and the mean expectation estimate associated with it was significantly lower than for the other quantifiers which communicated a similarly low amount (*not many, few, very few* and *only a few,* which appear to form statements of denial). Moxey and Sanford explain the results in terms of the expectancies signalled by these quantifiers. They argue that, because *a few* communicates a small amount but does not suggest that a higher amount was expected, the expectation estimates generated in response to materials quantified by this will be lower than the estimates associated with quantifiers which communicate a similar quantity (as indicated by the results from the first experiment) but which also signal that more would be expected.

Figure 3.1 Mean expectation estimates generated by participants in the Moxey and Sanford (1993b) study (Reproduced from Moxey and Sanford 1993a).



Mean estimates are collapsed across scenarios. Vertical lines indicate quantifier pairs which are reliably different at the $p<0.05$ level.

The evidence from this study would seem to provide initial support for a pre-condition of the Inference Model. The study clearly indicates that quantifiers which form negative statements may also signal an expectation violation. The next issue to be addressed concerns the impact of this aspect of presupposition denial upon the inferences generated in response to negatively quantified statements.

3.6.2 Content types and reference pattern

Further evidence which is consistent with the Inference Model derives from analysis of the content types generated in the Moxey and Sanford (1987) and Sanford et al (1996) continuation studies and, additionally, from the focus modifying effect associated with the type of connective conjoining quantified statement and continuation apparent in the Moxey and Sanford (1987) study.

Perhaps the main source of empirical support for the Inference Model originates in kinds of information typically related continuation tasks. Moxey and Sanford (1987) found that the continuations generated in response to a quantified sentence completion task (as described in Chapter 2) could be categorised in terms of four main types of content. The first type, which will be termed *Reason why (RW)* throughout this thesis, reflected continuations which related some type of explanation for the predicate of the quantified sentence being true. This is illustrated in (14). Here, the second sentence provides a reason for the action described in the quantified sentence.

(14) Many fans went to the match. They were committed supporters of the team.

The second category of continuation, termed *Reason why not (RWN)*, reflected continuations which provided an explanation for the predicate of the quantified sentence being false. This is illustrated in (15) by the continuation providing a reason for being absent from the match.

(15) Few fans went to the match. They had become bored with football and decided to get drunk instead.

The third category, termed *Consequence of the number (CONSEQ)*, related some consequence of the number of those for whom the predicate was true. In (16), the continuation relates a consequence of the lack of attendance at the match for those who went.

(16) Few fans went to the match. They felt unhappy that the stadium was nearly empty.

The final category was composed of continuations which did not fit with the other categories developed and was simply termed *Other*.

The continuations were categorised in terms of this classification scheme by six independent judges. The predominant continuation type associated with each of the quantifiers investigated was then assessed. Table 3.2 provides a summary of the main results obtained. The results indicated that the quantifiers which were associated with compset focus (*few* and *very few*) were also distinguishable in terms of the main content type produced in response to them. These quantifiers tend to lead to continuations which provide an explanation for the predicate not being true (i.e. RWN), while this category is not apparent in the quantifiers associated with refset focus.

Table 3.2 The predominant continuation and reference types for the *full-stop* condition obtained by Moxey and Sanford (1987)

Quantifier	*Continuation type*	*Reference type*
few	RWN	compset
very few	RWN	compset
a few	OTHER	refset
only a few	OTHER/RW	refset

The association of content type and quantifier was further investigated in two completion studies carried out by Sanford et al (1996). Table 3.3 shows the pattern of continuation types obtained in this study. These results again illustrate the relative predominance of RWN continuations for quantifiers associated with compset focus. Sanford et al (1996) suggest that this pattern of results may be indicative of the inferential activity associated with negatively quantified sentences and, consequently, of the mechanism which underlies compset focus. They argue that the presupposition denial signalled by the negative sentence engenders a *virtual question* concerning why *more x* was not the case. This virtual question then leads to a search for a plausible explanation for the expectation violation, which is typically answerable in terms of a property of the compset (hence the binding of the plural anaphor to the compset). So, given a negatively quantified sentence such as (17), the presupposition that someone believed that *many* fans would be expected to go to the match may lead to the virtual question: *Why did more fans not go?* The processor may then attempt to find a plausible answer for this question, typically in terms of some particular property of those whose behaviour did not fit the expectation (in this case, those who did not go to the match). Inferential activity may then be expected to

centre upon possible reasons for not going to the match (for example, because the entities which comprise the compset were doing something else instead).

(17) Not many of the fans went to the match. They...

Table 3.3 Percentage continuation types from Sanford et al (1996)

	Content Type			
Quantifier	*RW*	*RWN*	*CONSEQ*	*Other*
Not quite all	10	50	0	40
Not all	3.3	66.7	0	26.7
Less than half	10	53.3	0	33.3
Not many	3.3	63.3	0	33.3
Few	3.3	40	0	53.3
Not quite 10%	13.3	50	0	36.7
Not quite 30%	6.7	50	0	43.3
Not quite 50%	0	56.7	0	43.3
Not quite 70%	13.3	33.3	0	40
Not quite 90%	33.3	30	3.3	26.7
Nearly all	36.7	0	0	63.3
Almost all	50	0	0	50
More than half	43.3	0	3.3	53.3
Many	26.7	0	0	73.3
A few	26.7	0	0	70
Nearly 10%	43.3	0	3.3	53.3
Nearly 30%	40	0	0	56.7
Nearly 50%	40	3.3	3.3	53.3
Nearly 70%	40	0	0	60
Nearly 90%	33.3	0	0	66.7

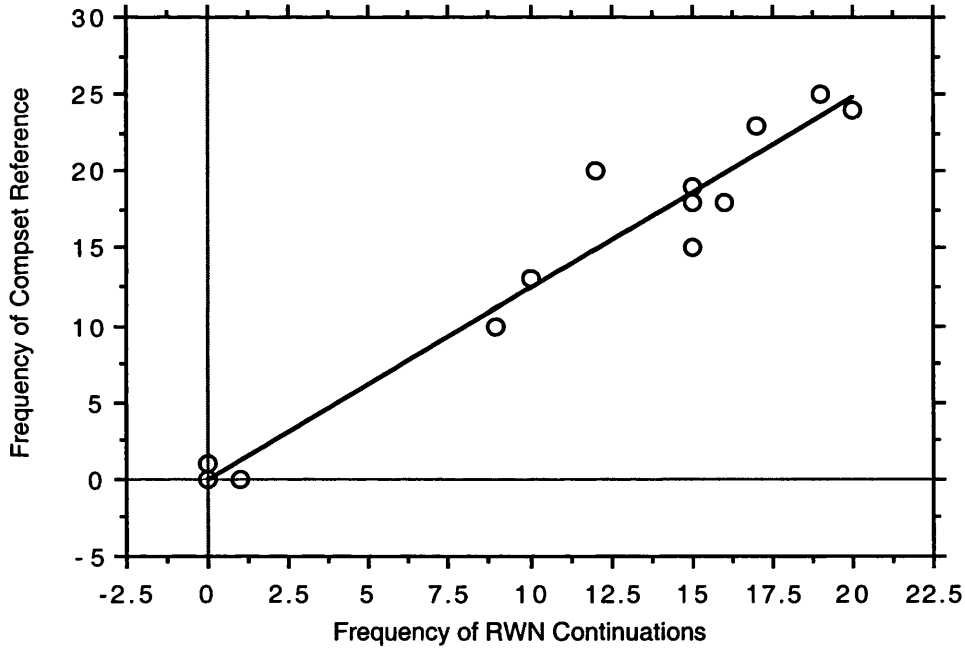
Analyses of the continuation type typically associated with compset focusing quantifiers (i.e. *Reasons Why Not*) provides a line of evidence which is consistent with the Inference Model. Negative quantifiers appear to trigger a particular type of inferential activity and this inferential activity appears to be anchored to some property of the compset.

3.6.2.1 Re-analysis of the Sanford et al (1996) sentence completion data

While the evidence related above demonstrates that both compset reference and RWN production are predominately associated with the same quantifiers, the extent to which the relative frequencies of these are correlated remains unclear. In order to further determine the relationship between compset reference and RWN production, a regression analysis

was carried out using the sentence completion data from Sanford et al (1996) Experiments 1 and 2. This produced a significant positive correlation between the frequencies of compset reference and RWN production ($r = 0.986$, d.f. = 18, $p<0.01$). Figure 3.2 shows the regression plot of compset reference on RWN production. This indicates a linear relationship between the variables (described by the function $y = 0.986x + 0.037$). This finding further reinforces the contention that there is a strong relationship between the incidence of compset reference and RWN production and, consequently, is consistent with the Inference account.

Figure 3.2 Regression plot for the frequency of compset reference on RWN production in the Sanford et al (1996) sentence completion studies



3.6.3 Connectives and set reference

This sub-section considers a third strand of empirical evidence which is consistent with the Inference Model. This evidence relates to the impact of connectives upon the pattern of focus and content type in quantified sentence completion tasks.

3.6.3.1 Connectives as cue-phrases for inferences

According to the Inference Model, the observed focal pattern associated with negatively quantified sentences results from the type of inferential field typically generated by such quantifiers. If this is the case, the focal pattern associated with a quantified statement should be sensitive to the addition of any linguistic device which signals a specific class of inference. That is, if the linguistic device constrains the pattern of inference, then the resulting reference pattern should alter as a function of this constrained inference. One such category of linguistic items is apparent in the case of connectives. These are a class of function words[1] which can signal the relationship between adjacent pieces of text (e.g. Halliday & Hasan, 1976; Millis & Just, 1994; Traxler, Bybee, & Pickering, 1997). In this sense, they have sometimes been referred to as *cue-phrases* (e.g. Knott, 1996).

Connectives are a complex and varied group of linguistic expressions which signal a diverse range of textual relationships (Knott, 1996). Three specific types of connective are of particular importance here as their impact upon the reference pattern associated with quantified sentences has been partially investigated by Moxey and Sanford (1987). These are additive, causal and adversative connectives, which are defined by Murray (1997) as follows. Additive connectives (e.g. *and*), signal a general elaboration upon preceding text. Causal connectives signal a cause and effect relationship between two pieces of text. These can be further cleaved into those which specify a cause-then-effect relationship (e.g. *and so*) and those which signal effect-then-cause relationships (e.g. *because*). Adversative connectives (e.g. *but*) indicate that the forthcoming text contrasts with, or in some way limits the scope of, the preceding text.

There is a considerable volume of evidence to suggest that the generation of inferences is often required in discourse comprehension (McKoon & Ratcliff, 1992; Potts, Keenan & Golding, 1988). This can be illustrated with reference to (18) below. Here, in order to

[1] In broad terms, function words are held to serve a syntactic function (e.g. Taylor & Taylor, 1990) and thus provide the processor with specific instructions concerning the way in which text should be processed.

fully understand the text, the reader may use existing world-knowledge (concerning the hazardous nature of ice), rather than information explicitly contained in the text, to infer that the situation described in the second sentence had caused the incident described in the first.

(18) John fell on the ground and broke his leg. The ice had made the pavement slippery.

Millis, Golding and Barker (1995) suggest that connectives may have a positive influence upon the probability that inferences are generated during the processing of text. They argue that, given a sentence such as (19) (Millis et al, 1995, p30), the presence of the connective *because* explicitly signals that the water spillage was caused by the little girl's actions. Millis et al suggest that the reader must then generate bridging inferences in order to represent the causal relationship between the sentence clauses (for example, that the vase contained water).

(19) Water spilled on the glass stand near the bookshelf because the little girl stuffed flowers into the vase.

Millis et al (1995) reported the results from three studies which investigated the impact of connective on inference generation. They concluded that while the presence of an additive connective (*and*) or a temporal connective (*after*) did not appear to have an impact upon the generation of inferences, the presence of a causal connective (*because*) was associated with the occurrence of causal inferences.

3.6.3.2 The impact of connective on focus pattern and content type

As briefly mentioned in Chapter 2, the Moxey and Sanford (1987) continuation study also investigated the effect of connective on the focus patterns associated with quantified statements. In this study, the basic quantified statement (see Table 2.3, Chapter 2) and

pronominal prompt were conjoined by one of four possible connectives (*and, but, because* or *full-stop*). Table 3.4 provides a summary of the reference patterns generated in the continuations.

The *full-stop* condition illustrates the familiar reference pattern: *few* and *very few* are associated primarily with compset focus, while *a few* and *only a few* seem to be refset focusing. This basic reference pattern for the compset focusing quantifiers, however, appears to be sensitive to the addition of other forms of connective prior to the pronominal prompt. Specifically, the connectives *and* and *but* seem to have the effect of (reliably) reducing the incidence of compset focus, while *because* appears to be associated with an amplification of compset focus. This focus altering effect of *because* is most notable in the case of the quantifier *only a few*. Here the strong refset focus associated with the other connective conditions is markedly changed to principally compset under the influence of *because*.

Table 3.4 Reference patterns and principal content types from Moxey and Sanford (1987)

Quantifier	*Connective*	*Refset*	*Compset*	*Content type*
few (x)	.	35	62.5	RWN
	and	95	0	OTHER/CONSEQ
	but	72.5	25	CONSEQ
	because	5	85	RWN
very few (x)	.	20	72.5	RWN
	and	82.5	10	OTHER
	but	70	25	CONSEQ
	because	0	90	RWN
a few (x)	.	95	0	OTHER
	and	100	0	OTHER
	but	100	0	OTHER
	because	90	5	RW
only a few (x)	.	95	5	OTHER/RW
	and	97.5	0	OTHER
	but	97.5	0	CONSEQ
	because	22.5	60	RWN

This evidence, then, would seem to suggest that the focus effect associated with quantifiers is, at least to some extent, sensitive to the presence of certain types of

connective. In particular, the connective *because* would seem to be associated with both an amplification of compset focus and an increase in the occurrence of the RWN type of continuation. This pattern of results would appear to be generally consistent with the Inference Model. That is, if content type is taken as an index of inferential activity, the results would seem to indicate that *because* may have the effect of triggering an increase in causal inferences (in line with the findings of Millis et al, 1995) which, in turn, lead to an increase in the rate of complement set references.

The evidence, however, would seem to be inconsistent with set-driven accounts of quantifier focus. For, if the basic compset focus effect is mediated by some type of semantic mechanism operating on quantifiers of a certain type, it would seem reasonable to suggest that focus should already be established by the time the connective is encountered (that is, if the quantifier itself initiates the focusing effect, focus should be established at this point). If this is the case, the connective would be expected to be interpreted in the context of the pre-existent focus and the initial focus pattern preserved. While, of course, it is conceivable that connectives may, under some circumstances, initiate some refocusing mechanism (that is, alter a previously established focus), no mechanism has yet been advanced to explain this. The fact that the pattern observed is not consistent with the set-driven models would seem to suggest that the mechanism which gives rise to the focus effect is more complex than is suggested by the current semantic perspectives.

3.7 Chapter Summary and Conclusions

The chapter began by describing the semantic explanations for complement set focus proposed by Kibble (1997) and Devlin (1997). It was then argued that both these explanations were insufficient to account for the phenomenon of compset focus. Specifically, the account proposed by Kibble appeared to be inadequate for two principal reasons. First, the account suggests that downwards monotonicity might be the trigger for compset focus. However, because the available empirical evidence suggests that this

property is not a unique predictor of the occurrence of the phenomenon (compset focus has also been shown to occur in the presence of nonmonotone and M↑ quantifiers) it would appear that the proposal is not supported by the available evidence. Secondly, Kibble's model does not suggest a mechanism which would account for the variable rates of complement set focus associated with different quantifiers. The model proposed by Devlin (1997) suggests that some property of negativity may be the causal antecedent of compset focus. However, negativity is underspecified and, like Kibble's model, Devlin does not account for the variable rates of compset focus. The chapter then introduced the Inference-Model. This was argued to present a potentially more complete explanation for compset focus in that:

1. It is motivated by an understanding of negativity gleaned from a considerable amount of empirical data concerning the representation and function of negativity in language.
2. It is supported by evidence from the content types present in continuations.
3. Connectives are thought to exert an impact upon inference generation and there is some evidence to suggest that they also impact upon the reference pattern present in the completions.
4. Judgements concerning the property of S-negation, which is also associated with presupposition denial, associated with different quantifiers may provide a basis for predicting the variable rates of compset focus. That is, quantifiers differ in their acceptability in tests of S-negation and this variability in sentence acceptability may correlate with compset focus rates associated with the quantifier. If this is the case, then this may constitute additional evidence that presupposition denial triggers complement set focus.

Section B

Off-line Studies of the Mechanism Underlying Complement Set Focus

The chapters in this section report the results from four sentence completion experiments.

Chapter Four

The Impact of Denial on Complement Set Reference in Free and Constrained Sentence Completion Tasks.

4.1 Introduction

The primary objective of this chapter is to provide an empirical evaluation of the central argument expressed by the Inference Model. That is, that compset reference arises as a consequence of denial. The Inference Model suggests that when a reader interprets a quantified expression as denying some presupposition, this expectation violation leads the processor to search for an explanation for why the expected state of affairs did not occur. This, it is argued, then puts processing focus upon the sub-set which is associated with the expectation violation (i.e. the compset) and thus renders the compset available for subsequent pronominal reference.

The claim that it is denial which triggers compset reference is not, however, a straightforward matter to evaluate. For, as demonstrated in Chapter 3, the properties of downwards monotonicity and denial are often associated with the same quantifier. The main purpose of this chapter, then, is to determine whether it is the property of denial, rather than that of downwards monotonicity, which is implicated in the phenomenon of complement set reference. The chapter reports the results from a series of studies which address this and other critical issues.

4.2 Experiment One:

The Effect of Denial in an Unconstrained Sentence Completion Task

4.2.1 Introduction

The evidence from previous sentence completion studies, discussed in Chapters 2 and 3, strongly suggests that both referential focus and content types generated in sentence completions vary as a function of quantifier type, however, the nature of the pronominal probe task itself may be problematic. For, while naturally occurring examples of complement set focus have been observed (e.g. Kibble, 1997), there is a lack of evidence concerning their frequency of incidence and their distribution as a function of quantifier type. It is therefore possible that both the observed focus effect and the variance in content type occur as a consequence of focus being forced by the pronominal probe. That is, when the processor encounters the pronoun it is forced to search for a plausible textual antecedent, thereby generating the observed focus effect. In this case, the phenomenon would be driven not simply by quantifier type, but also by task demands.

Experiment 1 investigates the patterns of focus and content type generated in response to a wide range of quantifiers (including monotone decreasing quantifiers which produce affirmation and denial statements). Specifically, it tests the following hypotheses: first, that the focal and content patterns observed will reproduce those found in work using pronominal probes. Second, that monotone decreasing quantifiers which produce denials will be associated with compset reference while those which produce affirmations will not. Third, that denials will be associated with the production of RWN type content.

4.2.2 Pre-test of denial / affirmation

One of the central objectives of Experiment 1 is to test a claim which arises from the Inference Model. That is, that it is the property of denial, and not that of downwards monotonicity, which triggers the phenomenon of complement set reference. As discussed

in Chapters 2 and 3, the tag tests which act as diagnostics of denial and affirmation depend upon subjective judgements of grammaticality. It was therefore necessary to carry out a pre-test of denial/affirmation on the quantifiers used in the experiment.

4.2.2.1 Method

Materials and design

The pre-test used the three of basic tag-tests described in Chapters two and three. These are presented in Table 4.1. Each of the three test sentences were combined with the eight quantifiers used in Experiment 1, creating a total of 24 test items. The test items were presented in a random order.

Table 4.1 Pre-test test sentences

Qx of the men were happy, and neither/so is Mary.
Qx of the students like maths, do/don't they?
Qx of the men liked the food, and **Qx** of the women did either/too.

Where **Qx** was one of: *A few, Nearly all, At least 10%, No less than 10%, Few, Not quite all, At most 10%* or *At most 90%.*

The tag options are separated by the slashes. In each case the first tag option indicates denials and the second affirmations.

Participants

Fifteen postgraduate students and members of academic staff took part in the pre-test. All were naive to the purpose of the study.

Procedure

Each participant was given a printed sheet of the pre-test items and, for each item, asked to tick the tag option which they felt was most grammatically acceptable.

4.2.2.2 Results

For each quantifier, the total numbers of denials or affirmations ticked were summed over the tests and participants. These scores were then used as an operational index of the extent to which each quantifier could be considered a denial or an affirmation. The results are presented in Table 4.2. In the table, the *Index of Denial Strength* refers to the number of denial tags checked expressed as a proportion of the total number of tags checked for that quantifier (i.e. the number of denial tags divided by 45).

Table 4.2 Total number of denial and affirmation tags checked in the pre-test for Experiment 1

Monotonicity	Quantifier	Denials	Affirmations	Index of Denial Strength
Monotone Increasing	A few (x)	2	43	0.04
	Nearly all (x)	2	43	0.04
	No less than 10% (x)	18	27	0.40
	At least 10% (x)	2	43	0.04
Monotone Decreasing	Few (x)	41	4	0.91
	Not quite all (x)	32	13	0.71
	At most 10% (x)	5	40	0.11
	At most 90% (x)	3	42	0.06

Two aspects of the data are of particular importance. First, for the monotone decreasing set, *few (x)* and *not quite all (x)* predominately formed denials, while *at most 10% (x)* and *at most 90% (x)* were predominately affirmative. Second, for the monotone increasing set, all the quantifiers appeared to be strongly affirmative with the exception of *no less than 10%* (for which a total of 18 denial tags were checked).

4.2.3 Method

Materials and design

The experiment employed the same 30 basic scenarios used in Experiments 1 and 2 of the Sanford et al (1996) study. In one of these materials, the noun phrase was altered from *new medical school* to *new doctors* to render the material less ambiguous. These materials

were used because they cover a broad sample of possible situations. An example material is given in Table 4.3 and a full list of the basic materials is provided in appendix 1. Each material consisted of a quantified sentence followed by a space for the completion.

Table 4.3 Example material from Experiment 1

Qx of the job applicants did well in the interviews...
Where: **Qx** was one of: *A few, Nearly all, At least 10%, No less than 10%, Few, Not quite all, At most 10%* or *At most 90%.*

The two sets of quantifiers (comprised of 4 monotone increasing and 4 monotone decreasing quantifiers) used in the pre-test constituted the group of quantifiers used in this experiment. These quantifier sets allowed for comparisons to be made between the effects of monotone decreasing and increasing sets and also between the effects of relative denial strength within the sets.

The quantifiers were each combined with the 30 scenario sentences. This created a total of 240 individual experimental sentences. To avoid contrast effects, each participant was given only one sentence to complete (and therefore saw only one quantifier / material combination). This created a completely independent group design.

Participants

240 undergraduate students volunteered to take part in this study. None had taken part in a similar experiment and all were naive to the purpose of the study.

Procedure

Each participant was given a single quantified sentence printed on paper and asked to produce two sentences as a sensible completion of the text. They were encouraged to

perform the task as spontaneously as possible. There was no further task required of the participant.

4.2.4 Results

The main purpose of this experiment was to investigate the extent to which the patterns of reference and content type observed in previous research are maintained in the context of a free continuation task. In terms of the Inference Model, the principal argument is that the content of the continuations should reflect the inferences generated in response to the quantified sentence and, therefore, explain the patterns of reference typically associated with the quantifier.

In this experiment, the continuations were analysed in terms of reference pattern and content types generated. It should be noted, however, that because there was no explicit requirement for the continuations to make reference to one of the sets associated with the quantifiers (since there was no explicit pronominal probe), it was not necessarily the case that any occurrence of set reference would occur.

Scoring procedure for the reference patterns

Unlike many of the previous experiments reviewed, where participants were forced to use a pronoun, it was not practical to ask participants to specify the referents of any pronouns that they may have used. However, previous research of this type (e.g. Sanford et al, 1996) has shown that there is generally good agreement between subjects and judges regarding the intended referent. For this study, two independent judges[1] classified the continuation responses as appearing to refer to the reference set, the complement set, as generalisations, or to other referents. Where a claim is made that complement set reference occurred, this was consonant with the judges' intuitions.

[1] The author of this thesis was one of the judges.

Reference pattern analysis

The analyses were confined to the first sentence of the continuation produced by the participants. This was because the content of later sentences is likely to have been constrained by the participants' own initial statement, in addition to experimental item itself. The continuations produced included ones where anaphoric references (including pronominal and other forms of reference) were made to both the reference set and the complement set. The types of anaphoric reference observed included pronouns, full noun phrases and contrastive references (e.g. *the others*). The continuations were analysed for references to the sets deriving from the quantifiers, and the results are presented in Table 4.4.

Table 4.4 Frequencies of compset, refset, general set and other references in Experiment 1 continuations (pronominal reference in brackets). Maximum score = 30

Quantifier	Denial Strength	Compset	Refset	Gen/ All	No reference	Total
A few (x)	0.04	9 (0)	14 (9)	4 (0)	3	30 (9)
Nearly all (x)	0.04	9 (1)	12 (10)	6 (4)	3	30 (15)
No less than 10% (x)	0.40	14 (3)	11 (8)	2 (0)	3	30 (11)
At least 10% (x)	0.04	8 (0)	14 (9)	5 (0)	3	30 (9)
Total M↑		40 (4)	51 (36)	17 (4)	12	120 (44)
Few (x)	0.91	18 (10)	4 (3)	3 (2)	5	30 (15)
Not quite all (x)	0.71	19 (9)	6 (1)	5 (3)	0	30 (13)
At most 10% (x)	0.11	16 (1)	6 (6)	2 (1)	6	30 (8)
At most 90% (x)	0.06	14 (1)	3 (1)	4 (1)	9	30 (3)
Total M↓		67 (21)	19 (11)	14 (7)	20	120 (39)
Grand Total		107 (25)	70 (47)	31 (11)	32	240 (83)

The reference patterns observed were analysed in terms of both pronominal reference (as this can be used as an index of focus) and general anaphoric reference (i.e. including all categories of anaphoric reference to some part of the quantified set).

Pronominal reference pattern

Focus is primarily associated with pronominal reference, so this will be considered first. Comparison of the pooled reference data for the monotone decreasing set with that of the increasing set indicated a reliably higher incidence of pronominal reference to the compset 17.5 % versus 3.3%, $\chi^2(1) = 11.56$, $p < 0.001$) for the monotone decreasing set. This set was also associated with a reliably lower incidence of refset reference (9.2% versus 30%, $\chi^2(1) = 13.29$, $p < 0.001$). This pattern strikingly confirms earlier observations, and indicates that the earlier findings were not simply due to reference being forced through the use of a pronominal probe.

Most significantly for the Inference Model, within the decreasing set, the denial associated quantifier set was associated with a reliably higher incidence of compset reference than the affirmation set (31.6% versus 3.3%, $\chi^2(1) = 13.76$, $p < 0.001$).

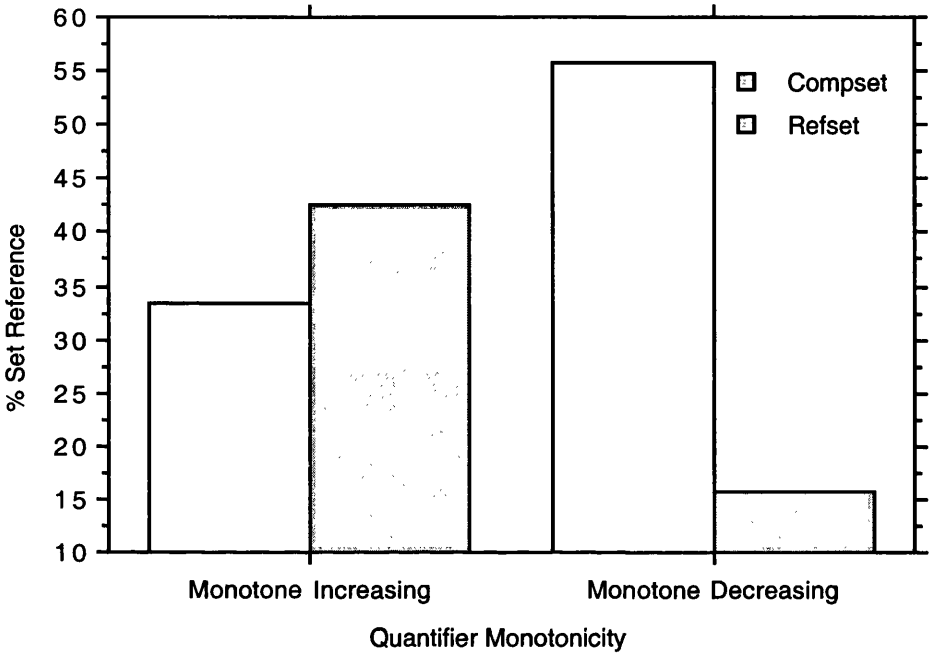
Three crucially important conclusions can be drawn from these findings. First, the results demonstrate the reality of spontaneous pronominal compset reference. Secondly, pronominal compset references are predominately associated with downwards monotonicity. Thirdly, pronominal compset references are largely restricted to denials.

General anaphoric reference pattern

Figure 4.1 shows the frequency of compset and refset reference associated with the monotone increasing and monotone decreasing sets of quantifiers (the frequencies are pooled for each quantifier set and expressed as a percentage of the total number of responses for that set). The monotone decreasing set was associated with a reliably higher incidence of compset reference (55.8% versus 33.3%, $\chi^2(1) = 6.80$, $p < 0.01$) and a reliably lower incidence of refset reference (15.8% versus 42.5%, $\chi^2(1) = 14.62$, $p < 0.001$). This general pattern is also consistent with that observed in previous work using pronominal probes.

Within the monotone decreasing set, comparison of pooled references for the denial associated quantifiers with those of the affirmative set yielded no reliable difference for either compset reference (χ^2 (1) = 0.73, ns) or refset reference (χ^2 (1) = 0.05, ns).

Figure 4.1 Percentage general anaphoric reference to compset and refset for the monotone increasing and monotone decreasing quantifier sets



Scoring procedure for the content types

Two independent judges[2] classified the content of the continuations according to the classification scheme devised by Moxey and Sanford (1987; Sanford et al, 1996). The categories, and appropriate tests, are related below.

Reason Why Not (RWN): As described in Chapter 3, this type of continuation provides an explanation for the falsity of the predicate. That is, it explains why the predicate is not true for some subset of the quantified set.

Test: Ask whether <predicate false> BECAUSE <continuation> is a plausible assertion.

[2] The author of this thesis was one of the judges.

Reason Why (RW): Again, as related in Chapter 3, this type of continuation provides an explanation for the truth of the predicate.

Test: Ask whether <predicate true> BECAUSE <continuation> is a plausible assertion.

Other: This category was used to classify the continuation when neither of those above were appropriate.

Content type analyses

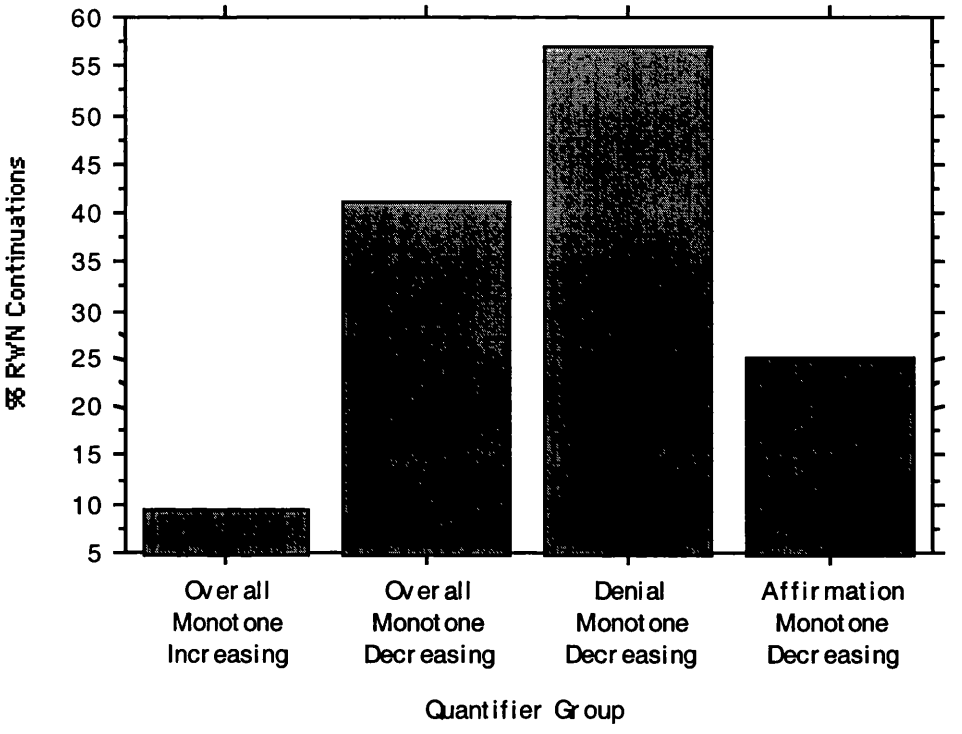
Again, the analyses were based upon the content of the first sentence. There was 86% agreement between the judges. After discussion there was only one instance where classification was not possible. The frequencies of the content types are shown in Table 4.5.

Table 4.5 Frequencies of continuation types for Experiment 1

Quantifier	RW	RWN	Other	Not clear	Total
A few (x)	5	2	23	0	30
Nearly all (x)	7	2	21	0	30
No less than 10% (x)	8	3	19	0	30
At least 10% (x)	14	4	12	0	30
Total M↑	34	11	75	0	120
Few (x)	2	18	9	1	30
Not quite all (x)	1	16	13	0	30
At most 10% (x)	5	8	17	0	30
At most 90% (x)	1	7	22	0	30
Total M↓	9	49	61	1	120
Grand Total	43	60	136	1	240

Figure 4.2 shows the overall frequency of RWN responses which were associated with each of the four critical quantifier groupings (i.e. monotone increasing set, monotone decreasing set, denial monotone decreasing and affirmation monotone decreasing).

Figure 4.2 Percentage RWN continuations for the four quantifier groups. The frequency data was pooled for each group and then expressed as a percentage of the total number of responses for that group.



An overall comparison of the monotone decreasing and monotone increasing sets showed that RWN responses accounted for 41% of all responses for the monotone decreasing set and only 9% of the monotone increasing responses. A chi-square test indicated that the difference was reliable (χ^2 (1) = 24.06, $p < 0.001$). The reverse pattern was observed for RW responses, with those accounting for 7.5% of all responses for the monotone decreasing set and 28% of those for the monotone increasing set. Again, the difference was reliable (χ^2 (1) = 14.53, $p < 0.001$). These results are consistent with previous observations, again indicating that monotone decreasing expressions are associated with the production of RWN responses, while these are relatively rare for monotone increasing expressions.

Comparison of the denial and affirmation groups within the monotone decreasing set indicated that 56.6% of all responses for the denial set (composed of the quantifiers *few (x)* and *not quite all (x)*) were accounted for by RWN continuations, while only 25% of the affirmative set (*at most 10% (x)* and *at most 90% (x)*) responses were associated with RWN continuations. A chi-square test indicated that this difference was reliable (χ^2 (1) =

7.36, $p < 0.01$). This supports the hypothesis that denials induce a higher frequency of RWN responses than affirmations.

Analysis of the relationship between compset reference and RWN production

In order to illustrate the relationship between compset reference and RWN production, two regression analyses were carried out, using the data from all eight quantifiers. There was a significant positive correlation between the frequencies of pronominal compset reference and RWN production ($r = 0.913$, d.f.=6, $p < 0.01$). Figure 4.3 shows the regression plot of pronominal compset reference on RWN production. This indicates a linear relationship between the variables (described by the function $y = 0.913x - 1.291$).

Figure 4.3 Regression plot for the frequency of pronominal compset reference on RWN production

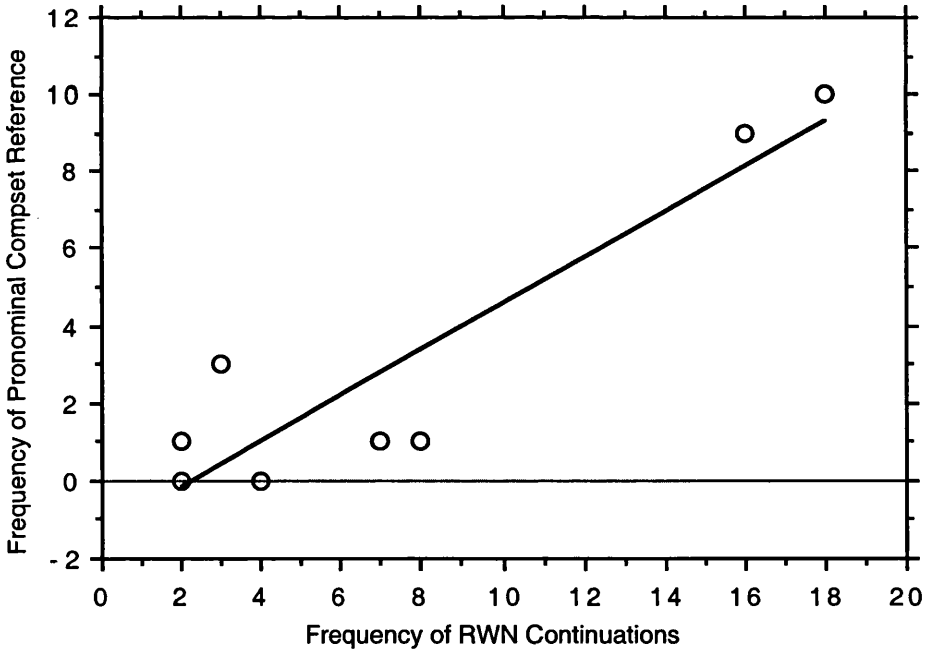
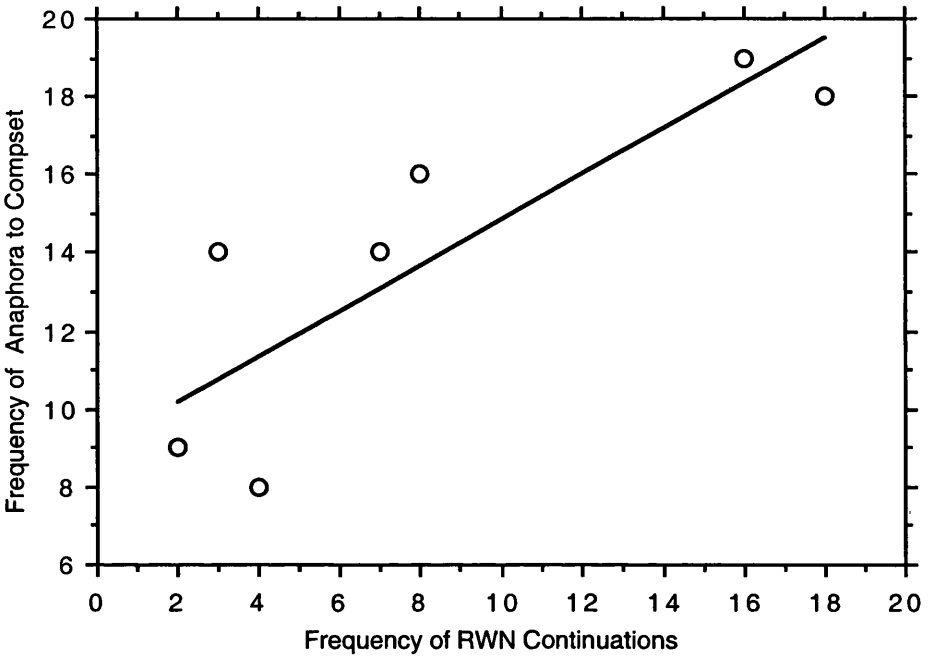


Figure 4.4 Regression plot for the frequency of general anaphora to compset on RWN production



There was also significant positive correlation between the frequencies of general complement set anaphora and RWN production ($r = 0.860$, d.f.=6, $p < 0.01$). Figure 4.4 shows the regression plot of general compset anaphora on RWN production. The plot indicates a linear relationship between general compset anaphora and RWN production described by the function $y = 8.986 + 0.86x$ (where y is the frequency of general compset reference and x is the frequency of RWN production). These results suggest a strong relationship between the incidence of both pronominal compset reference and general compset anaphora, and the production of RWN continuations. This relationship is consistent with the Inference Model.

4.2.5 Discussion

There were two principal aims behind this experiment. The first was to attempt to reproduce the focal and content patterns observed in previous work using pronominal probes. The second was to attempt to differentiate the effect of downwards monotonicity upon focus and content from that of denial. This second experimental aim is of considerable importance because while the Inference Model suggests that it is denial

which triggers compset focus (through initiating an inferential activity which results in the search for RWN), Kibble's set-driven account suggests that downwards monotonicity alone underlies the phenomenon. The experiment, therefore, provided an opportunity to empirically test the predictions made by both models.

Comparison of the responses generated for the monotone increasing set with those generated in response to the monotone decreasing set indicated that the decreasing set was associated with reliably higher incidences of both pronominal and general complement set anaphora and also of RWN content. This pattern reproduces that found by previous work and thus supports the first experimental hypothesis. In order to appreciate the full significance of the results relating to pronominal complement set reference, it is first necessary to distinguish the function performed by pronominal reference from that of other forms of anaphora. In this respect, pronouns are used to refer to currently focused textual entities (e.g. Gordon, Grosz & Gilliom, 1993; Gordon & Hendrick, 1998; Sanford, Moar & Garrod, 1988), while full noun phrases and other forms of non-pronominal reference may be used when the referent is not currently focused (Vonk, Hustinx & Simons, 1992). Thus, the data relating to pronominal compset reference obtained in this experiment provide direct evidence that the focal patterns observed in previous work (using pronominal probes) are not an artefact of the experimental methodology employed in those, but rather reflect a genuine phenomenon.

Within the monotone decreasing set, analysis indicated that the denial associated quantifier set produced a reliably higher incidence of both pronominal compset reference and RWN content than did the affirmative set. This finding supports the hypothesis that denial expressions would be associated with a higher incidence of compset focus than affirmative expressions, and also the hypothesis that denials would be associated with a higher incidence of RWN content. This evidence is consistent with the Inference Model but is not explicable within the current semantic framework proposed by Kibble.

The Inference Model is also supported by the observation of significant positive correlations between both pronominal compset reference and general compset anaphora, and the production of RWN continuations. Since the Inference Model suggests that it is the search for RWN which may trigger compset reference, the model is supported by this direct evidence of a relationship between the variables.

A potentially problematic aspect of the data is apparent in the case of the monotone increasing quantifier *no less than 10% (x)*. This quantifier had the highest rate of compset reference (including three occurrences of pronominal reference) for the monotone increasing set. However, while the quantifier is associated with a relatively high denial strength (0.4), it would appear to specifically deny the expectation that less than 10% may have been the case. Thus, in terms of the Inference Model, it should trigger a search for reasons why more was the case and consequently lead to refset continuations. This finding would, therefore, appear initially hard to explain in terms of either the Inference Model or Kibble's account and is further investigated by Experiment 2.

4.2.6 Experiment 1 auxiliary study

4.2.6.1 Introduction

A major question which remains to be resolved concerns the observation that different monotone decreasing quantifiers are associated with different rates of compset reference. In Chapter 3, it was suggested that this variable rate may be attributable to a variability in the ease with which specific quantifiers can be held to assert a denial. The main experimental conditions of Experiment 1 provided some evidence which was consistent with this hypothesis in that these indicated that monotone decreasing quantifiers which strongly express denial are associated with pronominal compset reference while those which are strongly affirmative do not. However, since Experiment 1 only considered monotone decreasing quantifiers which formed either clear affirmations or clear denials, it was not possible to determine whether or not the effect exists as a continuum. In order to

address this question and allow the overall effect of denial strength on compset rate to be better assessed, two additional conditions were run subsequent to the main experiment (employing the analytically monotone decreasing *less than 10% (x)* and its analytically monotone increasing counterpart *more than 10% (x)*).

The inclusion of *less than 10% (x)* was motivated by two factors. First, previous research (Moxey & Sanford, 1993a; Sanford et al, 1996) has indicated that the quantifier *less than x (x)* is associated with a middle-ranking rate of compset reference. Secondly, this quantifier appears to yield intuitively ambiguous results in tests for denial (see Chapter 3, section 3.5.2) and, therefore, may also be associated with a middle-ranking denial strength.

4.2.6.2 Pre-test of denial / affirmation

A pre-test was carried out in order to provide an operational index of denial strength for the quantifiers used in the auxiliary conditions. The method employed was identical in most respects to that reported for the pre-test for Experiment 1. The sole difference was that only two quantifiers (*less than 10% (x)* and *more than 10% (x)*) were used and, consequently, each participant responded to a total of six test items. The results are presented in Table 4.6. The most notable feature of these is that *less than 10% (x)* is associated with a denial strength of 0.4. This finding is consistent with the intuitive judgements made in Chapter 3 and suggests that the quantifier forms weaker denials than *few (x)* and *not quite all (x)*, but stronger denials than *at most x (x)*. If denial strength is related to pronominal compset reference, this result would suggest that *less than 10% (x)* should be associated with a similarly middle-ranking rate of compset reference.

Table 4.6 Total number of denial and affirmation tags checked in the pre-test for the auxiliary study.

Monotonicity	Quantifier	Denials	Affirmations	Index of Denial Strength
M↑	More than 10% (x)	3	42	0.06
M↓	Less than 10% (x)	18	27	0.4

4.2.6.3 Method

Materials and Design

The quantifiers *less than 10% (x)* and *more than 10% (x)* were each combined with the 30 basic scenario sentences used in the main study. This created a total of 60 test items.

Participants

60 undergraduate students participated in the study. All were unaware of the purpose of the study and had not taken part in a similar experiment.

Procedure

This was identical to the main study.

4.2.6.4 Results and discussion

The reference patterns and content types were analysed by two independent judges[3] according to the classification schemes used in Experiment 1. There was a 93% initial agreement between the judges for reference type and 90% initial agreement for content type. In both categories, differences were easily resolved upon discussion. The results are presented in Tables 4.7 and 4.8.

Table 4.7 Frequencies of compset, refset, general set and other references in the auxiliary study (pronominal reference in brackets). Maximum score = 30

Quantifier	Mono-tonicity	Denial Strength	Compset	Refset	Gen/ all	No ref	Total
More than 10% (x)	M↑	0.06	7 (0)	11 (6)	1 (1)	11	30 (7)
Less than 10% (x)	M↓	0.4	11 (6)	5 (1)	3 (0)	11	30 (7)
Total			18 (6)	16 (7)	4 (1)	22	60 (14)

[3] The author of this thesis was one of the judges.

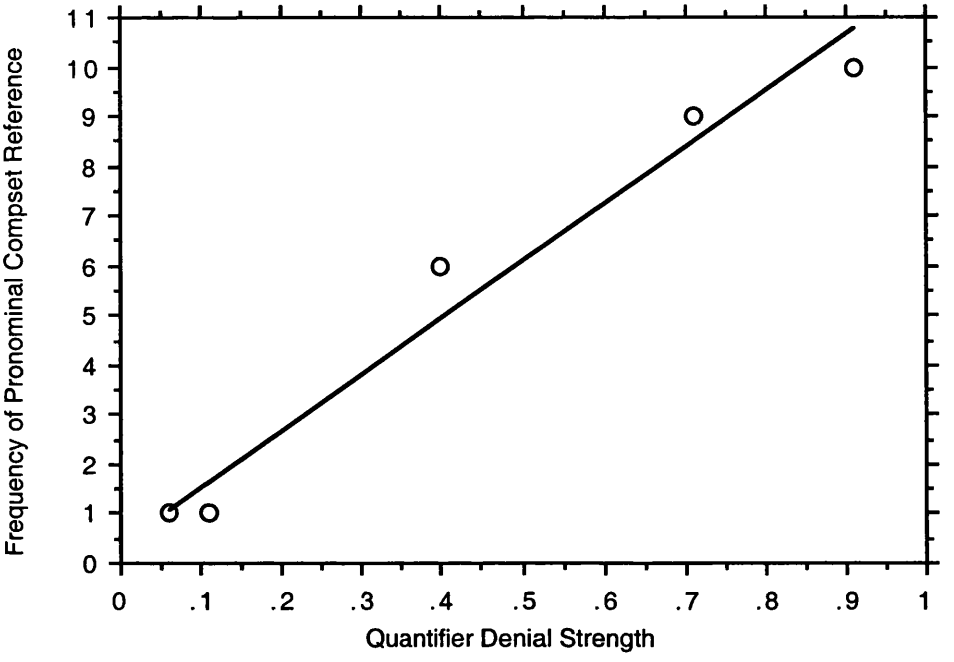
Table 4.8 Frequencies of continuation types for the auxiliary study

Quantifier	RW	RWN	Other	Total
More than 10%	9	0	21	30
Less than 10%	3	12	15	30
Total	12	12	36	60

Analysis of the relationship between the frequency of compset reference and quantifier denial strength

In order to investigate the overall relationship between quantifier denial strength (as indexed by the pre-tests) and the frequency of pronominal compset reference, a regression analysis was conducted using the data from the monotone decreasing quantifiers used in both the main and auxiliary studies. There was a significant positive correlation between the frequencies of pronominal compset reference and quantifier denial strength ($r = 0.984$, d.f.=3, $p < 0.01$). Figure 4.5 shows the regression plot of pronominal compset reference on quantifier denial strength. This indicates a linear relationship between the variables (described by the function $y = 0.984x - 0.424$). This would suggest that compset rate may be sensitive to the strength of denial communicated by the quantifier. This finding is of considerable importance because it provides some support for the claim that the observed variability in compset rate between quantifiers may be attributable to differences in the extent to which quantifiers signal denial.

Figure 4.5 Regression Plot for the frequency of pronominal compset reference on quantifier denial strength for the monotone decreasing quantifiers used in Experiment 1 and the auxiliary study.



4.3 Experiment Two:

The Effect of Denial in Forced Reference Continuations

4.3.1 Introduction

A key finding of Experiment 1, was that the property of denial (in conjunction with that of downwards monotonicity) is highly associated with the phenomenon of compset focus. However, the nature of the unconstrained task meant that participants were not necessarily required to refer to an antecedent and it is possible that the effect of denial may be demonstrated more clearly in a task which forces pronominal reference. A central aim of this experiment is to confirm, and clarify, the relationship between denial and compset focus by utilising the forced reference methodology adopted by previous research. The experiment tests the hypotheses that monotone decreasing quantifiers which form denials will be associated with a higher incidence of both compset reference and RWN content than monotone decreasing quantifiers which form affirmations. Since denial strength has already been operationally indexed for three of the four quantifiers used in the study, it

was not practical to conduct a denial/affirmation pre-test for the single quantifier *no more than X (x).* However, substitution of *no more than 10%* into the test sentences used previously would seem to indicate that the quantifier forms denials. This is demonstrated by the intuitive acceptability of sentences (1), (3) and (5) and the apparent unacceptability of (2), (4) and (6).

(1) *No more than 10%* of the men were happy, and **neither** is Mary. √

(2) *No more than 10%* of the men were happy, and **so** is Mary. ×

(3) *No more than 10%* of the students like maths, **do** they? √

(4) *No more than 10%* of the students like maths, **don't** they? ×

(5) *No more than 10%* of the men liked the food, and *no more than 10%* of the women did **either**. √

(6) *No more than 10%* of the men liked the food, and *no more than 10%* of the women did **too**. ×

Additionally, the experiment addresses two further issues of importance. First, it tests a key prediction which arises from Kibble's semantic model of complement set reference. Recall that Kibble suggests that compset reference cannot occur for non-proportional quantifiers since these cannot be internally negated (see Chapter 3 for a fuller description of the argument). This experiment, then, formally tests Kibble's hypothesis that compset reference will occur for proportional quantifiers and not for quantifiers which are non-proportional. Second, the experiment attempts to confirm the curious reference pattern observed for the monotone increasing quantifier *no less than 10% (x)* in Experiment 1. Since the quantifier seems to deny the expectation that less than 10% would have been expected, the Inference Model would predict that the quantifier should lead to refset continuations. However, Experiment 1 indicated that the quantifier was actually associated with a relatively high rate of compset reference.

4.3.2 Method

Design and materials

The experiment employed the first 20 of the 30 basic materials which were used in Experiment 1. Each material consisted of a quantified sentence followed by a plural pronoun (*they*). An example material is given in Table 4.9. Table 4.10 relates the quantifiers used in the experiment and summarises their main properties. Each of the quantifiers appeared in one of two forms: numerical (e.g. *at most 10 (x)*) or proportional (e.g. *at most 10% (x)*). This manipulation facilitates the testing of Kibble's claim that complement set reference will only occur for proportional quantifiers.

Table 4.9 Example material from Experiment 2

Qx of the football fans went to the game. They....
Where: **Qx** was one of: *At least 10/10%, No less than 10/10%, At most 10/10%* or *No more than 10/10%* (the slashes separate the numerical and proportional forms).

Table 4.10 Properties of the quantifiers used in Experiment 2

Quantifier	Monotonicity	Denial/ Affirmation
At least 10 (x)	M↑	affirmation
At least 10% (x)	M↑	affirmation
No less than 10 (x)	M↑	denial
No less than 10% (x)	M↑	denial
At most 10 (x)	M↓	affirmation
At most 10% (x)	M↓	affirmation
No more than 10 (x)	M↓	denial
No more than 10% (x)	M↓	denial

The quantifiers were each combined with the 20 scenario sentences, creating a total of 160 individual experimental sentences. To avoid contrast effects, each participant was given only one sentence to complete (and therefore saw only one quantifier / material combination). This created a completely independent group design.

Participants

160 undergraduate students volunteered to take part in this study. None had taken part in a similar experiment and all were naive to the purpose of the study.

Procedure

Each participant was given a single sentence printed on paper and asked to write down a sensible completion of the text (beginning after the pronoun *they*). Upon completing the sentence, they were asked to turn over the sheet and indicate the referent of the pronoun by ticking a category corresponding to one of: refset, the set in general, compset, all the set or "other". For the example given in Table 4.9, the potential choices were:

(a) The fans who went to the game.
(b) The fans in general.
(c) The fans who did not go to the game.
(d) All of the fans.
(e) Other.

If the "other" category was chosen, the participant was asked to specify the referent. Note that category (a) corresponds to the refset; (c) to the compset, and (b) and (d) represent set generalisations.

4.3.3 Results

The continuations were analysed in terms of reference pattern and content types generated in the experiment.

Reference patterns

The frequencies of the various reference categories checked by participants are presented in Table 4.11. There was no reliable difference between quantifiers which expressed a numerical amount and those expressing a proportional amount for compset (χ^2 (1) = 0.4, ns), refset (χ^2 (1) = 0.08, ns) or generalised reference (χ^2 (1) = 0.09, ns). This division was therefore collapsed for the remaining reference pattern analyses.

The finding that compset reference occurs for both proportional and non-proportional quantifiers is important because it is inconsistent with the semantic account proposed by Kibble.

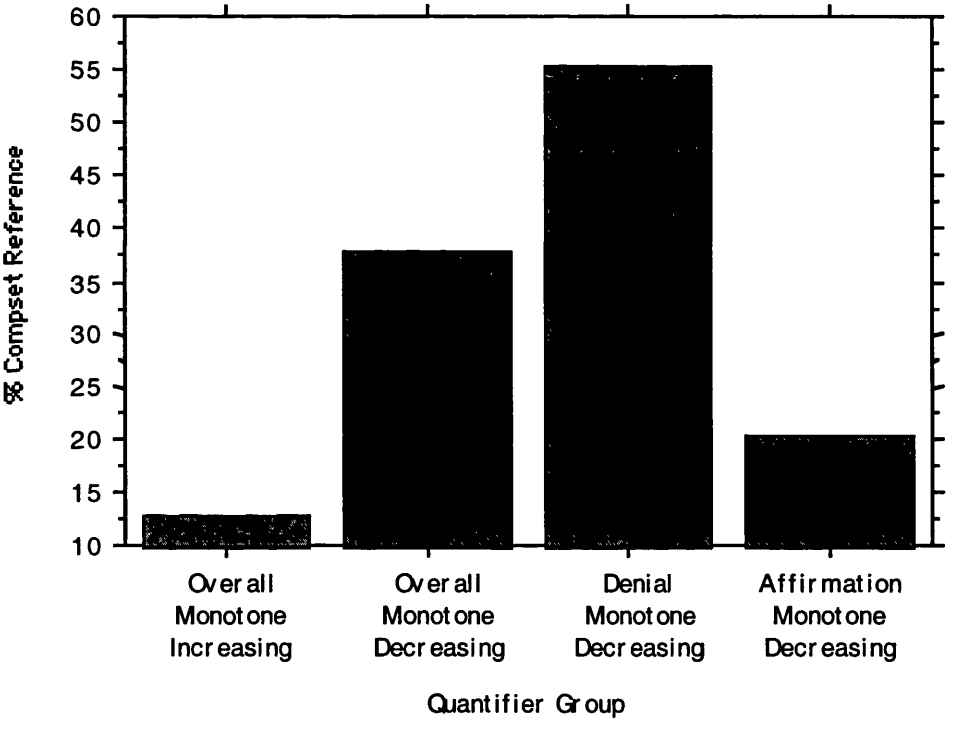
Table 4.11 Frequencies of compset, refset and general set references in Experiment 2 continuations (pronominal reference in brackets). Maximum score = 20

Quantifier	Denial/ Affirmation	Compset	Refset	Gen/ all	Total
At least 10 (x)	affirmation	0	20	0	20
At least 10% (x)	affirmation	1	19	0	20
Total		1	39	0	40
No less than 10 (x)	denial	3	16	1	20
No less than 10% (x)	denial	6	11	3	20
Total		9	27	4	40
At most 10 (x)	affirmation	3	16	1	20
At most 10% (x)	affirmation	5	15	0	20
Total		8	31	1	40
No more than 10 (x)	denial	12	4	4	20
No more than 10% (x)	denial	10	8	2	20
Total		22	12	6	40

Figure 4.6 shows the overall frequency of compset references which were associated with each of the four critical quantifier groupings (i.e. monotone increasing set, monotone decreasing set, denial monotone decreasing and affirmation monotone decreasing). Overall, there was a reliably higher frequency of compset reference associated with the monotone decreasing set of quantifiers (χ^2 (1) = 10, $p < 0.005$). This is consistent with the pattern observed in previous research.

Within the monotone decreasing set, the denial expressing *no more than x (x)* was associated with a reliably higher incidence of compset reference than was the affirmative *at most x (x)* (χ^2 (1) = 6.5, $p < 0.025$). This finding is consistent with the hypothesis that denials lead to compset references. Additionally, 22.5% of responses for the monotone increasing denial expression *no less than x (x)* referred to the complement set. This rate was reliably higher than for the affirmative monotone increasing *at least x (x)* (χ^2 (1) = 6.4, $p < 0.025$), but reliably lower than the monotone decreasing/denial *no more than x (x)* (χ^2 (1) = 5.45, $p < 0.025$). This result confirms the observation made in Experiment 1.

Figure 4.6 Percentage compset reference for the four quantifier groups in Experiment 2. The frequency data was pooled for each group and then expressed as a percentage of the total number of responses for that group.



Content type patterns

Two independent judges classified the content of the continuations according to the classification scheme used in Experiment 1. There was 95% initial agreement between the judges. The remaining instances were easily resolved upon discussion. The frequencies of the various content types generated are shown in Table 4.12.

There was no reliable difference between numerical and proportional quantifiers for RW (χ^2 (1) = 0), RWN (χ^2 (1) = 0.32, ns) or other continuations (χ^2 (1) = 0.44, ns).

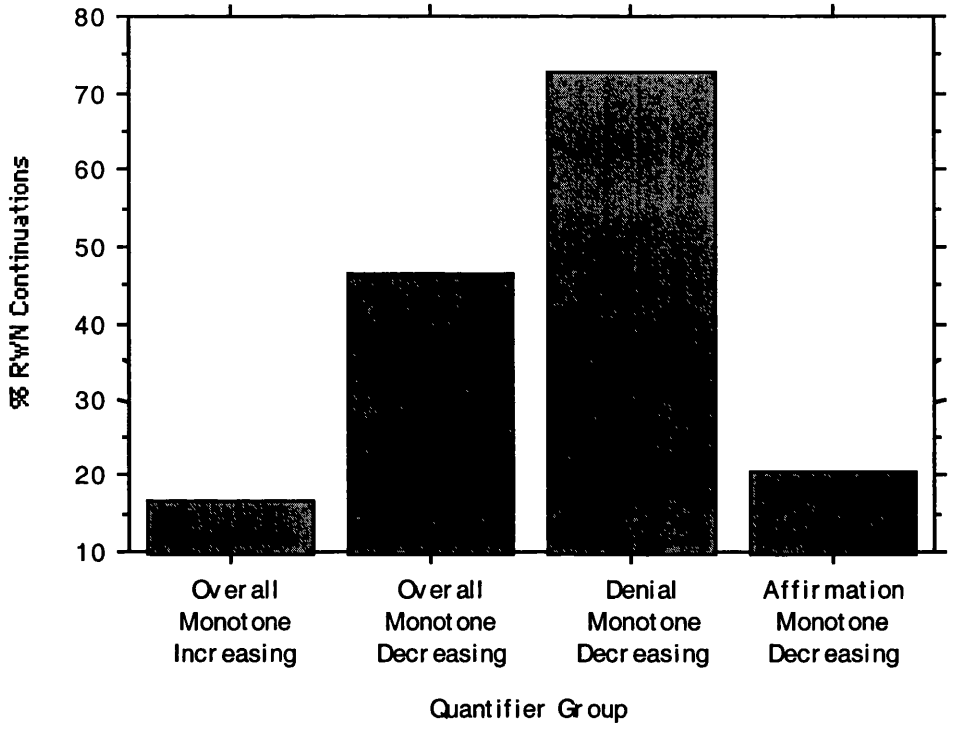
Table 4.12 Frequencies of the content types in Experiment 2 continuations

Quantifier	RW	RWN	Other	Total
At least 10 (x)	13	1	6	20
At least 10% (x)	11	0	9	20
Total	24	1	15	40
No less than 10 (x)	10	6	4	20
No less than 10% (x)	11	6	3	20
Total	21	12	7	40
At most 10 (x)	12	3	5	20
At most 10% (x)	11	5	4	20
Total	23	8	9	40
No more than 10 (x)	2	13	5	20
No more than 10% (x)	4	16	0	20
Total	6	29	5	40

Figure 4.7 shows the overall frequency RWN responses which were associated with each of the four critical quantifier groupings. RWN continuations were reliably higher for the monotone decreasing set than the monotone increasing set (χ^2 (1) = 11.52, $p < 0.005$). This pattern is consistent with previous research.

Within the monotone decreasing set, the denial expressing *no more than x (x)* was associated with a reliably higher incidence of RWN than was the affirmative *at most x (x)* (χ^2 (1) = 11.91, $p < 0.005$). This finding repeats that obtained in Experiment 1 and supports the hypothesis that monotone decreasing quantifiers which form denials lead to RWN continuations.

Figure 4.7 Percentage RWN continuations for the four quantifier groups in Experiment 2. The frequency data was pooled for each group and then expressed as a percentage of the total number of responses for that group.



4.3.4 Discussion

The main aim of Experiment 2 was to confirm the findings of Experiment 1 using a forced reference task. The results indicated that the frequencies of both compset reference and RWN content were reliably higher for monotone decreasing quantifiers which form denials than for those which form affirmations. Thus, both the primary hypotheses were supported. These findings suggest that denial may be a better predictor of complement set reference than downwards monotonicity and are consistent with the claim that compset reference arises as a consequence of the search for reasons why an expected state of affairs did not occur (this search being reflected in the production of RWN continuations). Thus, the experiment provides further support for the Inference Model.

The data from this experiment would appear to be inconsistent with Kibble's semantic account in two respects. First, evidence from this study suggests that it is the property of denial, rather than that of downwards monotonicity, which triggers complement set

reference. Secondly, the finding that compset reference is equally prevalent for both proportional and non-proportional quantifiers runs contrary to Kibble's prediction that compset reference would only be possible in the case of proportional quantifiers. These strands of evidence strongly suggest that the semantic model, as it is currently formulated, is not correct.

The finding that 22.5% of the continuations generated in response to the monotone increasing quantifier *no less than 10/10% (x)* referred to the compset confirms the result obtained in Experiment 1. This result is problematic because it does not, on the surface, appear to be explicable either in terms of the Inference Model or Kibble's semantic model. This matter is further addressed by Experiment 3.

4.4 Experiment Three:

Expectation Mismatch and Complement Set Reference

4.4.1 Introduction

An unexplained aspect of the data produced in Experiments 1 and 2 was the finding that the monotone increasing quantifier *no less than 10/10% (x)* was associated with a higher rate of compset reference than was expected. From the perspective of the Inference Model, *no less than 10/10% (x)* would appear to deny the expectation that less than 10/10% may have been the case and should, therefore, lead to refset reference (since the processor should search for reasons why more than was expected was the case). There would appear to be two possible explanations for this anomalous finding. First, it is possible that participants simply misinterpreted the quantifier (possibly treating it as *no more than x (x)*). Such an explanation would appear to be plausible in light of evidence which suggests that double negatives (such as *no less than*) are particularly hard to process (e.g. Wason & Reich, 1979). A second possible explanation may relate to a

mismatch between world knowledge and the small quantity expressed by *no less than 10/10% (x)*. This can be explained with reference to sentences (1) and (2) below.

(1) No less than 10 of the fans went to the match.

(2) No less than 10% of the fans went to the match.

In both cases, the number or proportion expressed represents a very small amount relative to the number who would typically be expected to attend a football match. However, as related above, the quantifier appears to imply that less would have been expected. Thus, there would appear to be a mismatch between the numerical expectation based on world knowledge and the small amount which, it is implied, is larger than expected. Under such circumstances, there would exist two alternative expectation violations to explain: one based upon world knowledge and the other upon the denial communicated by the quantifier. Given this situation, some participants may focus on the world knowledge based expectation violation and attempt to explain the small amount (and thus produce compset references). It is possible that this type of world knowledge mismatch occurs in some of the materials and this may be the basis for the unexpectedly high rate of compset reference observed for *no less than 10/10% (x)*.

Experiment 3 tests these alternative explanations. In this the quantifier *no less than 10/10% (x)* was replaced with *no less than 80/80% (x)* in order to reduce the world knowledge based mismatch effect. If the observed rate of compset reference in the earlier experiments is due to a world knowledge mismatch, the rate of compset reference should be reduced or eliminated in the current study. However, if the rate of compset reference is attributable to participants having difficulty in processing the double negative, changing the number/proportion should have no effect on compset reference.

4.4.2 Method

Design and Materials

The quantifiers *no less than 80% (x)* and *no less than 80 (x)* were each combined with the 20 basic stimulus materials used in Experiment 2. This created a total of 40 experimental items.

Participants

Forty undergraduate students participated in the study. None had taken part in a similar experiment and all were naive to the purpose of the study.

Procedure

Participants were given a single sentence printed on paper and asked to write down a completion of the text (beginning after the pronoun *they*). After completing the sentence, they were asked to turn over the sheet and indicate the referent of the pronoun.

4.4.3 Results and discussion

The frequencies of the various reference categories checked by participants are presented in Table 4.13. The most notable characteristic of the data is that the incidence of compset reference is almost eliminated (there was only a single occurrence in the 40 completions). A chi-square comparison of this value with that of the *no less than 10/10%* set from Experiment 2 indicates that the difference is reliable (χ^2 (1) = 6.4, $p < 0.025$). This finding would appear to support the world knowledge expectation mismatch hypothesis and suggests that the finding for *no less than 10/10% (x)* cannot be attributed to the double negative being misinterpreted by participants.

Table 4.13 Frequencies of compset, refset, general set and other references in Experiment 3. Maximum score = 20

Quantifier	Compset	Refset	Gen/ all	Total
No less than 80% (x)	0	13	7	20
No less than 80 (x)	1	16	3	20
Total	1	29	10	40

4.5 Summary and Conclusions

This chapter reported the results from three sentence completion studies. Experiment 1 used a free continuation methodology to investigate the occurrence of spontaneous compset reference and RWN production for expressions of denial and affirmation. The results of this study indicated that spontaneous occurrences of compset reference and RWN content did occur and that these occurrences were generally restricted to monotone decreasing quantifiers which formed denials. Further analysis of the relationship between denial and compset reference indicated that the frequency of complement set reference appeared to be closely related to the relative strength of denial communicated by the quantifier. These findings are generally consistent with the Inference Model, but are not explicable in terms of Kibble's semantic account.

Experiment 2 reverted to the forced reference task used in previous work and attempted to confirm the relationship between denial and compset reference/RWN production observed in Experiment 1. The results of this experiment confirmed the main findings of Experiment 1 and, furthermore, indicated that compset reference occurred for both proportional and numerical quantifiers (a finding which is inconsistent with Kibble's semantic account).

Experiment 3 was conducted in order to further investigate an unexplained finding of Experiments 1 and 2. The results of this experiment suggested that the occurrence of complement set reference for the monotone increasing quantifier *no less than 10/10%*

could be explained in terms of a conflict between expectations derived from world knowledge and expectation information communicated by the quantifier.

Chapter Five

The Effect of Connective on Reference Pattern and Content Types in Negative Polarity Quantifiers

Experiment 4

5.1 Introduction

One aspect of the empirical support for the Inference Model previously considered in Chapter 3 concerned the impact of connective upon set reference pattern. The purpose of this chapter is to provide a more extensive investigation of this issue.

The Inference Model proposes that complement set reference arises as a result of the inferential activity triggered by quantifiers which signal denial. If compset reference does in fact result from inferential activity, then it would seem reasonable to suggest that the focal pattern associated with compset focusing quantifiers should be sensitive to the imposition of an inferential constraint. In contrast to the Inference Model, Set-Driven accounts suggest that quantifier focus effects emerge from some type of direct mapping between quantifier and a logical subset and, therefore, reference pattern should not be sensitive to inferential constraint.

Previous research (Moxey & Sanford, 1987; 1993a) has demonstrated that different types of connective may exert an influence on both the content type and reference pattern generated in quantified sentence completion tasks. Recall that Moxey and Sanford (1987) found that, for the negative polarity quantifiers *few (x)*, *very few (x)* and *not many (x)*, compset reference was attenuated in the conditions containing the connectives *and* and *but* and enhanced in the *because* connective condition.

Additionally, *because* also altered the reference pattern associated with the quantifier *only a few (x)* from predominately refset to predominately compset. This finding provides some evidence which suggests that the inferential constraint imposed by connectives may impinge upon reference pattern. However, the study offered no explanation of the mechanism by which the observed compset suppression associated with *and* and *but,* for negative quantifiers, might arise. Moreover, as one of the primary content types associated with these conditions was classified in the non-specific *other* category (the other main content type being *consequence of number*), it is impossible to fully judge the extent to which the content types may be indicative of any specific class of inference (if a specific class of inference was evident in the continuations, this might suggest an explanatory mechanism for the altered pattern of focus). The experiment described in this chapter seeks to expand the investigation of the effect of connective on reference pattern and content type across a wider range of quantified contexts (the original study employed only two basic scenarios). Specifically, Experiment 4 tests the following hypotheses. First, that the incidence of both RWN type continuations and complement set reference will be amplified in the context of the causal connective *because*. Second, that the inclusion of both the additive connective *and* and the adversative connective *but* will be associated with a reduction of RWN type continuations and compset reference.

5.2 Method

Materials and Design

The experiment employed the same 30 basic scenarios used in Experiments 1 and 2 of this thesis. An example material is given in Table 5.1. Each material consisted of a quantified sentence, containing a connective, followed by a plural pronoun (*they*).

Table 5.1 Example material from Experiment 4

Qx of the holiday makers lay on the beach **[connective]** they....
Where: **Qx** was one of: *Few, Only a few* or *Not quite all.* **[connective]** was one of: *period* (i.e. "."), *and, but* or *because.*

The three quantifiers used in this study reflect several specific empirical interests and concerns. *Few (x)* and *only a few (x)* were both investigated in the original Moxey and Sanford (1987) connective study and their inclusion in this study allows for a direct comparison between the results obtained using the two, non-partitative, stimulus materials used by Moxey and Sanford and those employed in this study. Additionally, previous research (Moxey & Sanford, 1987; Sanford et al, 1996) has demonstrated a strong and very reliable compset licensing effect of *few* across different contexts. Therefore, any effect of connective on focal pattern should be clearly evident. *Only a few* is of specific interest in that it is a non-monotone quantifier and previous research (Moxey & Sanford, 1987; 1993a) has suggested that it is particularly sensitive to the effect of the connective *because* . *Not quite all* is included because, unlike many of the quantifiers previously investigated, the compset is small relative to the refset. Additionally, this quantifier is associated with unusually high compset reference (Sanford et al, 1996).

The connectives which conjoined the basic quantified sentence with the plural pronoun were chosen so as to signal a broad range of relationships between the initial (quantified) statement and the subsequent anaphoric reference completion and, additionally, to enable direct comparison with the original Moxey and Sanford (1987) study. The *period* connective signalled no specific relationship and was included to act as a baseline comparison point for the other connectives. The functions of the remaining connectives were outlined according to Murray's (1997) classification in Chapter 3. To recap, these are as follows. The additive connective *and* is held to signal a general

elaboration of the preceding text. The causal connective *because* signals a cause and effect relationship such that, the situation described in the primary statement should arise as a result of the information contained in the completion. Finally, the adversative connective *but* signals a paradoxical relationship between the initial statement and the completion. Here the completion statement should reflect some sort of expectation violation.

The experimental conditions were 3 (quantifier type) X 4 (connective type). To avoid contrast effects, each subject was given only one sentence to complete (and therefore saw only one quantifier/connective/material combination). This created a completely independent group design.

Participants

360 first and second year undergraduate students participated in the experiment. All were naive as to the purpose of the experiment.

Procedure

Each participant was given a single sentence printed on paper and asked to write down a sensible completion of the text (beginning after the pronoun *they*). Upon completing the sentence, they were asked to turn over the sheet and indicate the referent of the pronoun by ticking a category corresponding to one of: refset, the set in general, compset, all the set or "other". For the example given in Table 5.1, the potential choices were:

(a) The holiday makers who lay on the beach.
(b) The holiday makers in general.
(c) The holiday makers who did not lie on the beach.
(d) All of the holiday makers.

(e) Other.

If the "other" category was chosen, the participant was asked to specify the referent.

5.3 Results

The continuations were analysed in terms of reference pattern and content types generated in the experiment.

Reference patterns

The frequencies of the various reference categories checked by participants are presented in Table 5.2 (note that the *gen / all* category relates the pooled frequencies for both the *set in general* and *all the set* categories. While these categories are not necessarily the same, both represent set generalisations).

Table 5.2 Frequencies of refset, compset, general set and other referents of the pronoun "they" in Experiment 4 continuations. Maximum score = 30

Quantifier	Connective	Refset	Compset	Gen / all	Other
Few (x)	.	4	22	3	1
	and	15	9	6	0
	but	8	12	10	0
	because	3	16	9	2
Only a few (x)	.	21	8	1	0
	and	20	4	6	0
	but	19	3	8	0
	because	11	13	5	1
Not quite all (x)	.	0	25	5	0
	and	3	26	1	0
	but	4	13	13	0
	because	0	29	0	1

A series of chi-square tests were used to analyse the summarised reference type frequencies associated with both the quantifiers and connective types in the experiment.

There was a global effect of quantifier for both compset (χ^2 (2) = 35.2, $p < 0.005$) and refset (χ^2 (2) = 58.4, $p < 0.001$) but not for general set references (χ^2 (2) = 2.18, $p > 0.25$). Figure 5.1 shows the overall frequency of complement set references for the three quantifier conditions (pooled over the four connective conditions and expressed as a percentage of the total number of responses for each quantifier). The graph indicates that the highest compset frequencies were associated with the *not quite all* condition (77.5%), this being reliably higher (χ^2 (1) = 7.6, $p < 0.01$) than *few* (49.1%) which was, in turn, reliably higher (χ^2 (1) = 11, $p < 0.005$) than *only a few* (23.3%). This pattern of results is generally consistent with previous research (e.g. Moxey & Sanford, 1987).

Figure 5.1 Percentage complement set reference for the three quantifier conditions in Experiment 4

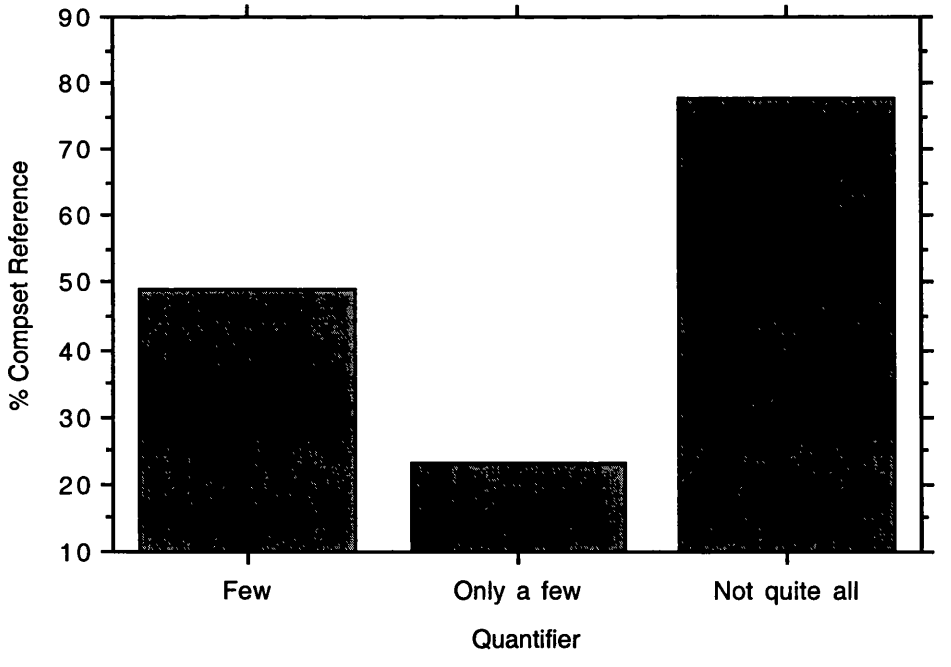
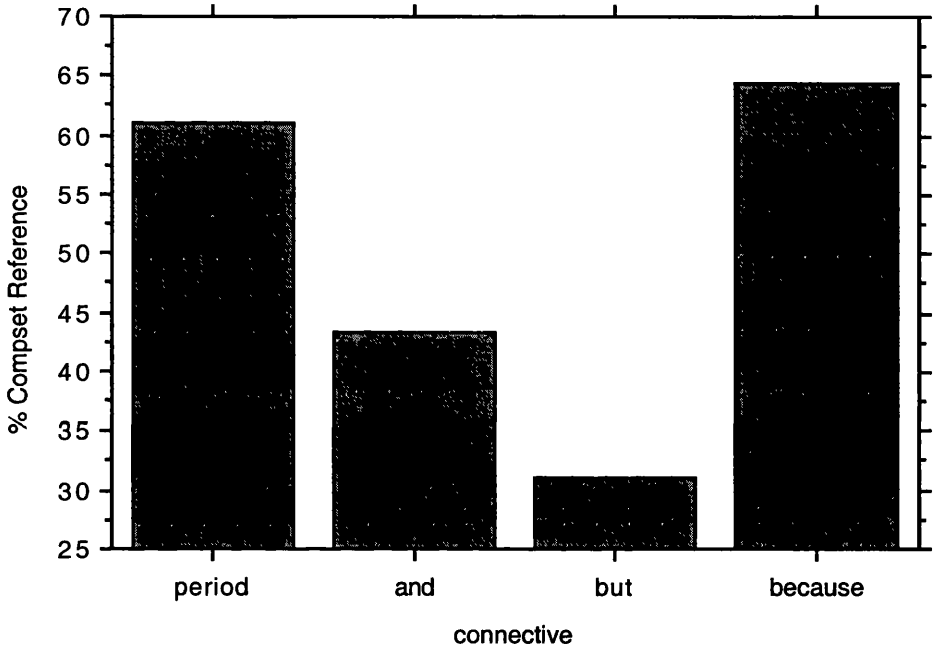


Figure 5.2 shows the overall frequency of complement set references associated with the four connective conditions (pooled over the three quantifier conditions and

expressed as percentage a of the total number of responses for each connective). The graph indicates that the highest compset frequencies are associated with the *because* (64.4%) and *period* (61.1%) conditions, with *and* being lower (43.3%) and *but* the lowest (31.1%). A chi-square test showed that there was a reliable global effect of connective on rate of compset reference (χ^2 (3) = 13.2, $p < 0.005$).

Figure 5.2 Percentage complement set reference for the four connective conditions in Experiment 4



The general numerical trend in this result would appear to partially parallel the results obtained by Moxey and Sanford (1987) in that *and* and *but* are again associated with a reduction in the rate of compset reference relative to the baseline *period* (although *because* does not appear to be associated with a marked increase in compsets). However, individual chi-square comparisons of *and*, *but* and *because* with the baseline *period* condition yielded a reliable difference for only the *but* versus *period* comparison (χ^2 (1) = 8.8, $p < 0.005$) (*and* versus *period*: being χ^2 (1) = 2.72, $p < 0.1$ and *because* versus *period*: being χ^2 (1) = 0.07, $p > 0.8$). Thus, the hypothesis that *because* would have the effect of amplifying compset reference is not supported by these findings.

Additionally, the predictions that *and* and *but* would be associated with a reduction in compset reference are only supported in the case of *but*.

Figure 5.3 shows the overall percentage of reference set references associated with the four connectives. Here *and* has the highest percentage of refset references (42.2%), followed by *but* (34.3%), *period* (27.7%) and *because* (15.5%). This general trend would seem to suggest that *and* and *but* may have the effect of amplifying reference set reference. However, while there was a global effect of connective (χ^2 (3) = 11.48, $p < 0.01$), none of the individual comparisons reached significance (*and* versus *period* χ^2 (1) = 0.16, $p > 0.5$; *but* versus *period* χ^2 (1) = 0.64, $p > 0.25$ and *because* versus *period* χ^2 (1) = 3.10, $p < 0.1$).

Figure 5.3 Percentage reference set reference for the four connective conditions in Experiment 4

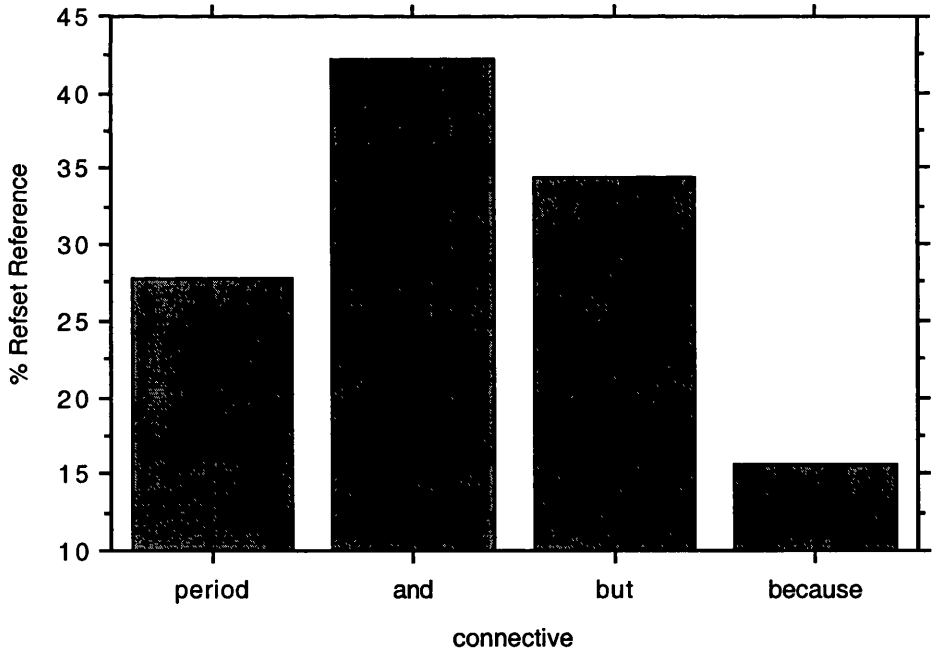
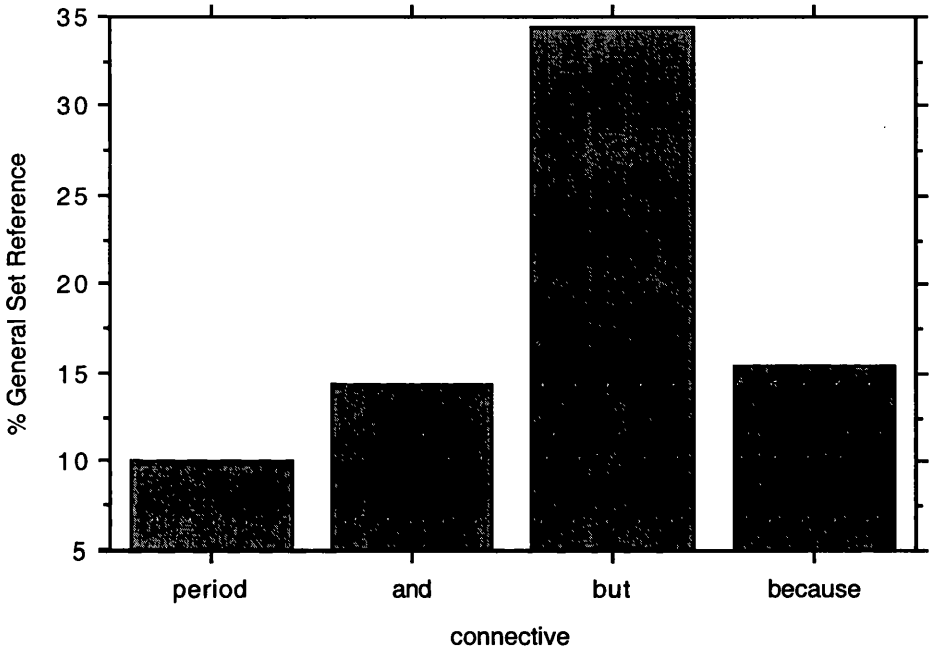


Figure 5.4 shows the overall percentage of general set references associated with the four connectives. Here *but* has the highest rate (34.4%), followed by *because* (15.5%), *and* (14.4%) and *period* (10%). Again, there was a global effect of connective (χ^2 (3) = 16.8, $p < 0.005$). Only one of the individual chi-square comparisons, that for *but*

versus *period*, reached significance (χ^2 (1) = 12.1, p < 0.005) (*and* versus *period* being χ^2 (1) = 0.72, p < 0.5 and *because* versus *period* χ^2 (1) = 1.08, p < 0.3). This would seem to suggest that *but* has the effect of amplifying set generalisations.

Figure 5.4 Percentage general set reference for the four connective conditions in Experiment 4



A further series of chi-square tests were used to analyse the effect of connective on the reference patterns for the individual quantifiers. For compsets, the only reliable effect of connective was found for *only a few* (χ^2 (3) = 8.85, p < 0.05) (neither *few* (χ^2 (3) = 6.42, p < 0.1) nor *not quite all* (χ^2 (3) = 6.39, p < 0.1) reaching significance). None of the individual comparisons between the *period* and the other connective conditions for *only a few* yielded a reliable result (*and* versus *period* being χ^2 (1) = 1.33, p < 0.3; *but* versus *period* χ^2 (1) = 2.27, p < 0.2 and *because* versus *period* χ^2 (1) = 1.19, p < 0.3).

For refsets, the only reliable effect of connective was for *few* (χ^2 (3) = 11.8, p < 0.01) (*only a few* did not yield a significant result (χ^2 (3) = 3.53, p < 0.2 and for *not quite all*

the expected frequencies were too low to allow the chi-square calculation). Within *few*, *and* was found to be associated with reliably more refset references than *period* (χ^2 (1) = 6.3, $p < 0.05$). None of the other individual comparisons were significant (*but* versus *period* χ^2 (1) = 1.33, $p < 0.3$ and *because* versus *period* χ^2 (1) = 0.14, $p < 0.7$).

For general set references, there were no reliable global effects for connective (*few* χ^2 (3) = 4.28, $p < 0.3$ and *only a few* χ^2 (3) = 5.2, $p < 0.2$ and *not quite all* had too low an expected frequency to allow calculation).

Classification of content type

Three independent judges[1] classified the content of the continuations according to the classification scheme described below. There was 83% initial agreement between the judges. The remaining instances were easily resolved upon discussion. The principal classification categories employed were those devised by Moxey and Sanford (1987). However, the complexity of the data obtained in this study necessitated the modification of one category (the *consequence of the small number* category used by Moxey and Sanford was subsumed into a more general category of *consequence*). Table 5.3 shows the content type frequencies obtained in the experiment.

Reason Why Not (RWN): As described previously, this type of continuation provides an explanation for the falsity of the predicate.

Reason Why (RW): Again as previously related, this type of continuation provides an explanation for the truth of the predicate.

Consequence (Consq) This category reflected continuations which related a plausible effect, or consequence, of the quantified statement. The category was comprised of the

[1] The author of this thesis was one of the judges.

following types of consequence:

(a) Consequence of Number: this represented the sole consequence category employed in both the Moxey and Sanford (1987) and Sanford et al (1996) studies. In this, the continuation relates a consequence of the number of people for whom the predicate is true.

E.g. *Few of the club members came to the AGM and they...did not enjoy themselves because the rest of the club weren't there.* (refset focus)

(b) Consequence of Predicate (true): This relates a consequence for the people for whom the predicate is true.

E.g. *Few of the soldiers were overweight and they...were too fat to fit through the tunnel on the assault course.* (refset focus)

(c) Consequence of Predicate (false): This relates a consequence for the people for whom the predicate is false

E.g. *Few of the patients liked the new hospital food and they...complained to the head cook.* (compset focus)

Contrary (Contra) This category reflected continuations which related a paradoxical or unexpected relationship between the initial-state and outcome reflected in the quantified statement and continuation. Continuations of this type often also included an explanation to account for the contrary consequence. The category was comprised of the following types of continuation:

(a) Contrary to Number: This type of continuation related a contrary consequence arising from the number of people whose behaviour fits the predicate.

E.g. *Few of the MPs were at the meeting but they...managed to discuss some of the important issues.* (refset focus).

(b) Contrary to Predicate (true): This relates a contrary consequence for the people for whom the predicate is true.

E.g. *Few of the club members came to the AGM but they...could not get in due to the door-staff being poorly informed.* (refset focus)

(c) Contrary to Predicate (false): Relates a contrary consequence for the people for whom the predicate is false

E.g. *Few of the holiday-makers lay on the beach but they...all still got a nice tan due to the good sun-shelter by the pool side.* (compset focus)

(d) In Spite of: Reflect continuations which imply that the situation described in quantified statement has occurred in spite of some property described in the continuation.

E.g. *Few of the accountants took up the retraining offer but they...thought it was a very good idea.* (compset focus)

Other: This category acted as a pool for any continuation which could not be classified in terms of the categories outlined above.

Content type patterns

Table 5.3 Frequencies of the content types in Experiment 4 continuations. Maximum score = 30

Quantifier	*Connective*	*RW*	*RWN*	*Consq*	*Contra*	*Other*
Few (x)	.	4	16	1	0	9
	and	2	1	20	0	7
	but	0	4	2	17	7
	because	1	29	0	0	0
Only a few (x)	.	17	9	0	1	3
	and	4	1	15	0	10
	but	2	0	13	7	8
	because	11	19	0	0	0
Not quite all (x)	.	0	24	0	0	6
	and	0	2	18	0	10
	but	0	1	5	8	16
	because	0	30	0	0	0

A series of chi-square tests were used to analyse the summarised content type frequencies associated with both the quantifiers and connective types in the experiment.

There was a global effect of quantifier upon both RWN type continuations (χ^2 (2) = 9.3, $p < 0.01$) and RW (χ^2 (2) = 47, $p < 0.005$); but not for consequence (χ^2 (2) = 0.7, ns) and contra (χ^2 (2) = 4.9, ns).

Figure 5.5 shows the overall frequency of RWN for the three quantifier conditions (again, pooled over the four connective conditions and expressed as a percentage of the total number of responses for each quantifier). The graph indicates that the highest RWN frequencies were associated with the *not quite all* condition (47.5%), then *few* (41.6%) and *only a few* (24.2%). Individual chi-square comparisons indicated that *not quite all* was reliably higher than *only a few* (χ^2 (1) = 9.11, $p < 0.005$) but not *few* (χ^2 (1) = 0.457, ns). In addition, *few* was reliably higher than *only a few* (χ^2 (1) = 5.58, $p < 0.02$).

Figure 5.5 Percentage RWN continuations for the three quantifier conditions in Experiment 4

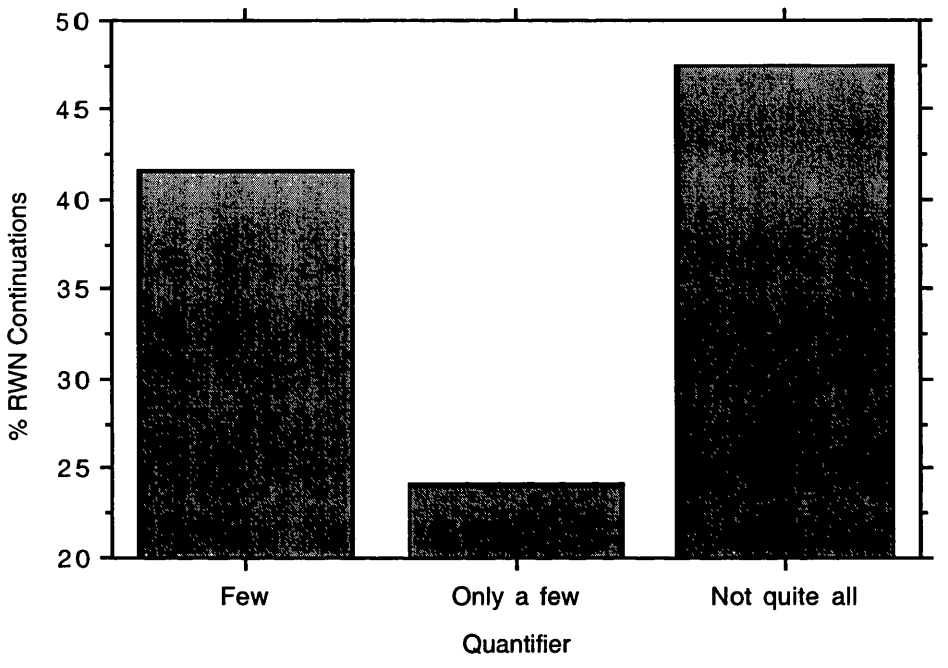


Figure 5.6 shows the overall frequency of RWN type continuations associated with the four connective conditions (again, the frequencies were pooled over the three quantifier conditions and expressed as a percentage of the total number of responses for each connective). The graph indicates that the highest incidence of RWN continuations are associated with the *because* (86.7%) and then *period* (54.4%) conditions; with, *and* and *but* being considerably lower (at 4.4% and 5.5% respectively). A chi-square test showed that there was a reliable global effect of connective on rate of RWN continuations (χ^2 (3) = 114.76 $p < 0.001$). Individual chi-square comparisons of *and*, *but* and *because* with the baseline *period* condition indicated that *and* and *but* were associated with reliably lower rates of RWN (with χ^2 (1) = 38.20, $p < 0.001$ for *and* and (χ^2 (1) = 35.85, $p < 0.001$ for *but*); while *because* was reliably higher (χ^2 (1) = 6.62, $p < 0.02$). The hypotheses that *because* would be associated with an increase in the incidence of RWN continuations, and that *and* and *but* would be associated with a decrease in these, are therefore supported.

Figure 5.6 Percentage RWN continuations for the four connective conditions in Experiment 4

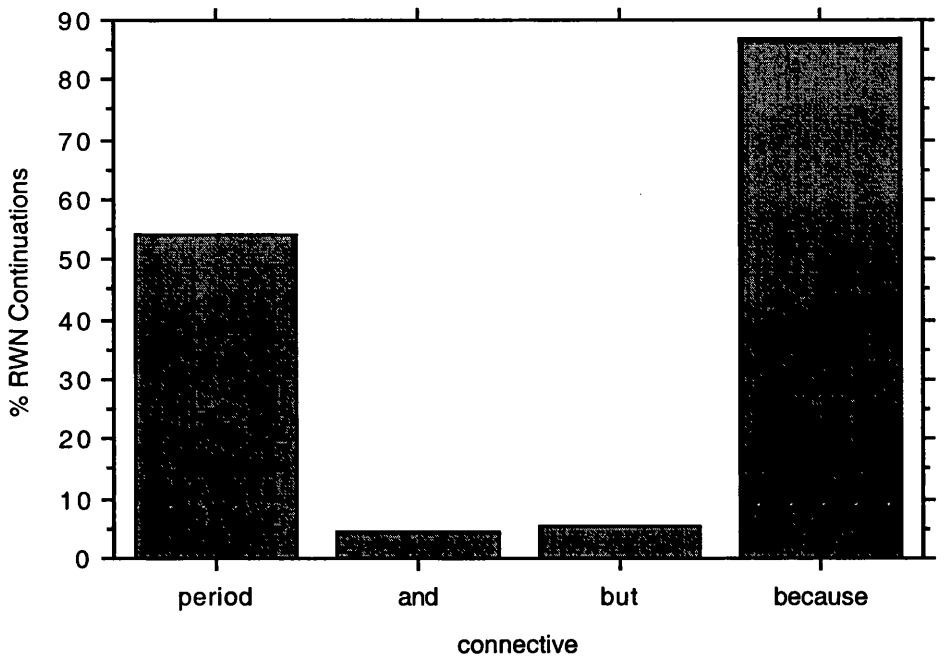


Figure 5.7 shows the overall frequency of consequence type continuations associated with the four connective conditions (again, the frequencies were pooled over the three quantifier conditions). The graph indicates that the highest incidence of consequence continuations are associated with the *and* (58.8%) and then *but* (22.2%) conditions; with, *period* and *because* being lower (at 1.1% and 0% respectively). A chi-square test showed that there was a reliable global effect of connective on rate of consequence continuations (χ^2 (3) = 99.5, $p < 0.001$). Individual chi-square comparisons of *and* and *but* (but not *because* as the expected frequency was too low to allow chi-square calculation) with the baseline *period* condition indicated that *and* and *but* were both associated with reliably higher rates of consequence continuations (with χ^2 (1) = 50.07, $p < 0.001$ for *and* and χ^2 (1) = 17.19, $p < 0.001$ for *but*).

Figure 5.7 Percentage consequence continuations for the four connective conditions in Experiment 4

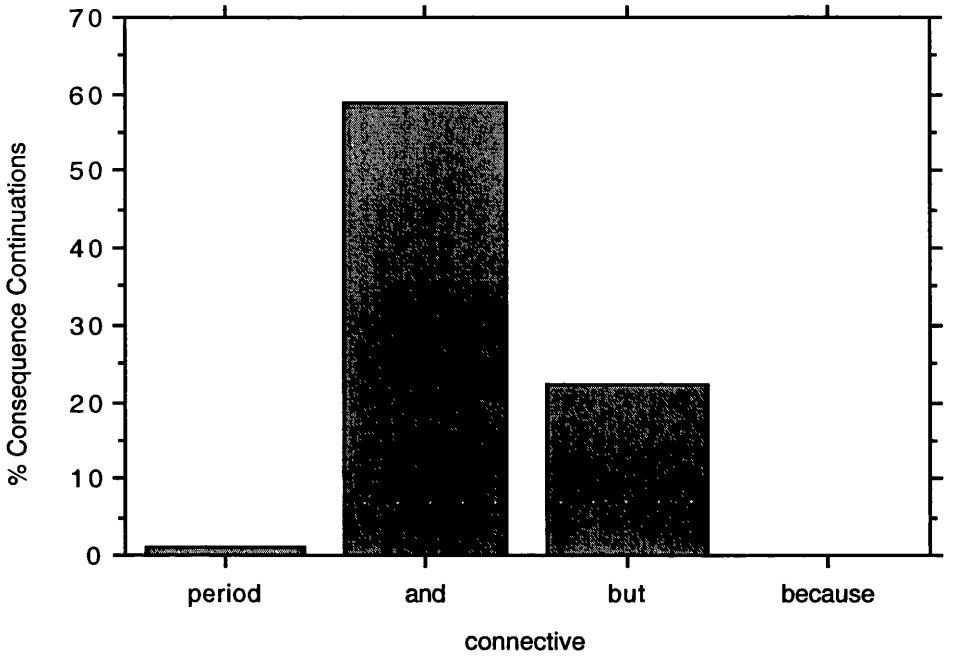
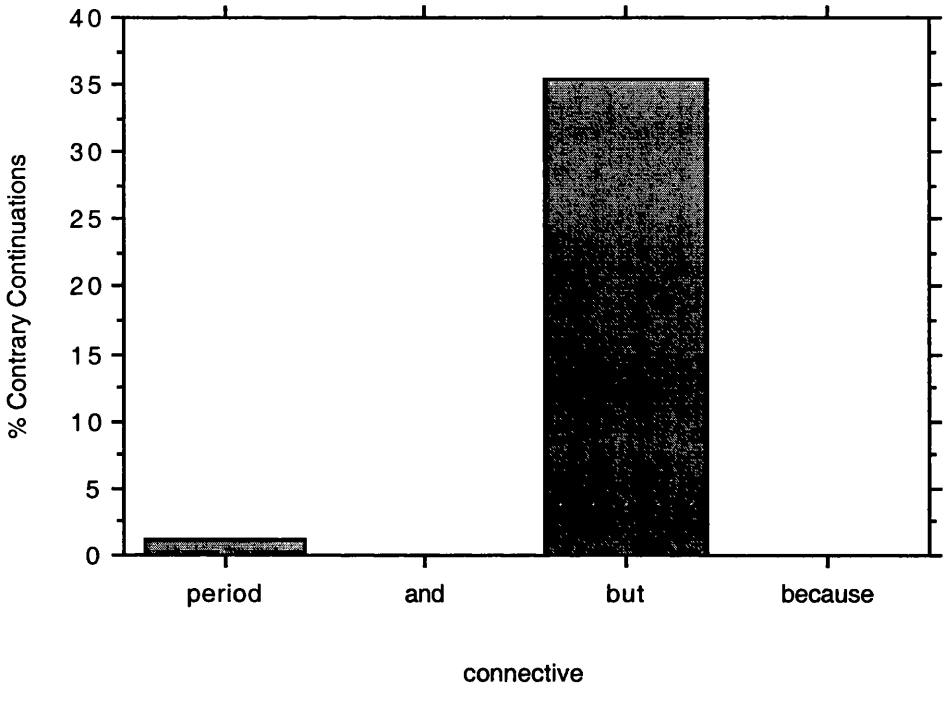


Figure 5.8 shows the overall frequency of contrary type continuations associated with the four connective conditions (again, the frequencies were pooled over the three quantifier conditions). The graph indicates that only the *but* condition was associated with an appreciable frequency of contra type continuations (35.5%), with *period* being associated with only 1.1% and no continuations of this type occurring for the *and* and *but* conditions. The relatively high incidence of contrary continuations associated with *but* would appear to be consistent with the suggested function of adversative connectives, that is that they signal that the associated text conflicts with, or runs contrary to, some expectation. There was no further analysis of these frequencies because the expected frequency was too low to allow chi-square calculation.

Figure 5.8 Percentage contrary continuations for the four connective conditions in Experiment 4



5.4 Discussion

In terms of the global effects observed for the quantifiers, the results of this study closely follow those reported in previous work (e.g. Sanford et al, 1996). Specifically, reliable differences were observed for the effect of quantifier on both frequency of compset reference and the incidence of RWN continuations, such that *not quite all (x)* and *few (x)* were associated with a higher incidence of compset reference and RWN continuations than *only a few*. This evidence is consistent with the central claim of the Inference Model that it is the search for RWN which may act as a trigger for compset reference.

The hypothesis that the causal connective *because* would be associated with an increase in compset reference relative to the baseline *period* connective was not supported by the experiment in terms of either the global effect of the connective or its specific effect upon the focus pattern of the non-monotone quantifier *only a few (x)*. While the fact

that these results are not consistent with the previous findings of Moxey and Sanford (1987) may be explicable in terms of the wider range of stimulus materials employed in the current study, an additional problem requires resolution. This second issue relates to the observation that although compset reference is not reliably amplified by *because*, the connective is associated with a reliable increase in RWN type continuations. Since the Inference Model suggests that the search for RWN may be the trigger for compset reference, the results of this study are problematic as the increased incidence of RWN for *because* do not appear to be associated with an increase in compset reference. A possible explanation for this observation may relate to a ceiling effect arising from two of the quantifiers used in this study. From the results it is apparent that the compset frequencies for the baseline (*period*) comparison point for the quantifiers *few* and *not quite all* are high (73.3% for *few* and 83.3% for *not quite all*) and this may have clouded the effect of the connective. That is, a high initial comparison point leaves less scope for amplification. A second possible explanation for this observation is discussed below. This issue would appear to be empirically resolvable by future research through the use of quantifiers which are associated with lower baseline compset rates.

The hypotheses that the adversative connective *but* would be associated with both a reduction in compset reference and a reduction in the incidence of RWN was strongly supported by the results (compset reference fell from 61.1% for *period* to 31.1% for *but*; while RWN fell from 54.4% for *period* to 5.5% for *but*). Interestingly, *but* was also associated with a reliable increase in general set reference (from 10% for *period* to 34.4% for *but*) and a small numerical (but not significant) increase in refset reference (from 27.7% for *period* to 34.3% for *but.*). This finding may be explicable in terms of the function of *but.*. It would seem reasonable to suggest that statements following the connective *but* may contrast with the inferences that might be expected to be true given the statement preceding the connective. Thus, given a statement such as (1),

(1) Few of the football fans went to the game.

some examples of plausible inferences might be that:

the fans, in general, were unmotivated;
the fans, in general, were disinterested in the outcome;
the game was not expected to be a good one;
the fans would be outnumbered by those of the opposition.

Introducing *but they...* as a continuation stem may then lead to denials of these plausible inferences, and such denials may tend to be based around either the fans in general or those who went to the match (i.e. the refset).

(2) ..but they had all bought season tickets.
(3) ..but they were all very happy when their team won.
(4) ..but they all enjoyed it very much.
(5) ..but they made a great impression with their singing and cheering.

The general observation that the connective *but* influences both reference pattern and content type may indicate that the focus pattern associated with negative quantifiers may be sensitive to the inferential constraints imposed by the connective and, thus, provides evidence which is consistent with the Inference Model. However, it is also important to note that while the occurrence of RWN continuations were almost eliminated in the *but* condition, the incidence of compset reference remained reasonably substantial. This would seem to suggest that the focal pattern associated with these negative polarity quantifiers does not solely result from the inferential constraints imposed by the connective *but*.

The hypothesis that the additive connective *and* would be associated with a decrease in compset reference was not supported by the results. For, while there was a numerical decrease in compset reference (from 61.1% for *period* to 43.3% for *and*), this was not

reliable. However, the results indicate that *and* was associated with both a decrease in RWN (from 54.4% for *period* to 4.4% for *and*) and an increase in consequence type continuations (from 1.1% for *period* to 58.8% for *and*). This observation would seem to indicate that the focus pattern associated with the quantifiers in this study is relatively insensitive to the variation in content generated in response to the connective *and.*

In contrast to previous work (e.g. Moxey & Sanford, 1987, 1993a; Sanford et al, 1996), this study has produced relatively little evidence of an association between complement set focus and the incidence of RWN continuations. In order to attempt an explanation of this inconsistency, it is first necessary to consider two possible mechanisms by which focus may be established.

The first possible mechanism proposes that the establishment of focus is deferred until the stimulus text has been fully processed. That is, focus is not established until the connective is encountered. From the perspective of the Inference Model, this mechanism would suggest that it is only at this point that inferences are generated and focus established upon the basis of these. If this is the case, it would be expected that there would be a high degree of association between the type of inferences reflected in the continuation and the referent of the pronoun. If, as previous research has suggested, RWN responses provide an index of the inferences which are typically associated with compset reference, it would seem reasonable to expect that this association would continue to hold. However, the evidence from this study would not appear to be consistent with this pattern.

The second mechanism suggests that focus may initially be established at an earlier point. In terms of the Inference Model, inferences would initially be generated in response to the implicit denial of expectation communicated by the quantifier and focus established upon this basis (i.e. before the connective is encountered). In this case, the subsequent inferential influence of the connective would either be integrated with the

already established focus (thus preserving this focus) or focus would have to be re-established. In either case, it would seem reasonable to assume that the content of the continuation would have to be grammatically consistent with the connective. That is, the content of the continuation may be constrained by the connective (e.g. RWN would be consistent with *because*, but not necessarily with *and*) and may not necessarily reflect the inferences which gave rise to the initial focus. This mechanism would seem to be consistent with the relative insensitivity of focal pattern to the presence of the connectives *and* and *because* observed in this study. Additionally, the diminished association between compset reference and the occurrence of RWN continuations may also be explicable within this framework. The effect of adversative connectives on focus pattern, and the time course of their integration within a quantified sentence, is investigated further by the on-line studies reported in Chapter 6.

Section C

On-line Studies of the Mechanism Underlying Complement Set Focus

Up until this point, this thesis has examined the mechanisms which underlie complement set focus in terms of off-line sentence completion studies. While these studies have produced evidence which is generally consistent with the Inference Model, a major question which remains to be resolved concerns the time at which focus is established. This question is of considerable importance because the two models considered in the preceding chapters give rise to differing predictions concerning the time course over which complement set focus is developed.

The set-driven accounts suggest that there exists some mechanism (triggered by quantifier downwards monotonicity or negativity) which allows a direct mapping between quantifier and sub-set, thereby making complement sets available for subsequent anaphoric reference. In terms of such models, focus should be established early and, consequently, the resolution of any subsequent pronominal anaphora should be immediate since the anaphor will directly bind to the focused set.

In contrast to the semantic accounts, the Inference Model suggests that complement set focus arises in a more complex way and, therefore, may not be established until a later point in comprehension. From the perspective of this model, there would appear to be two distinct possibilities (these were briefly discussed in Chapter 5). The first is that the inferences generated in response to a monotone decreasing/denial quantifier give rise to complement set focus during comprehension of the initial quantified statement itself. If this is the case, any subsequent anaphor should, again, simply bind to the focused set thereby allowing immediate anaphoric resolution. An alternative possibility is that focus is not developed until some point after an explicit anaphoric reference. This mechanism

would also seem to be theoretically plausible because monotone decreasing expressions appear to allow a range of potential referents (i.e. compset, refset and gen set continuations have been observed in continuation studies). Thus, comprehension of an anaphoric reference sentence may require that a fit is established between the content of the anaphoric sentence and the quantified sentence. Since the Inference Model proposes that compset focus is driven by inferences generated in response to denial, the resolution of compset anaphors may require that a fit is established between the virtual question raised by the denial and the content of the anaphoric sentence. If this is the case, the establishing of focus may be deferred until later in the anaphoric reference sentence.

To summarise, then, there are three possibilities for the timing of compset focus. The first is that focus is generated immediately upon encountering a monotone decreasing quantifier (by way of some semantic mechanism). In such a case, the resolution of subsequent anaphors should be immediate. Such a mechanism would be consistent with any set-driven account. The second is that focus is developed at a later point during comprehension of a quantified statement. Here, again, anaphoric resolution should be immediate. The third is that focus is not established until some point during the comprehension of an anaphoric reference sentence. In this case anaphoric resolution should be less immediate. Both the second and third possibilities would be consistent with the Inference Model.

At present, there exists only a small volume of empirical work which has any bearing upon this question. The evidence yielded by this work, however, would appear to be partially supportive of both early and later focus. For, while Experiment 4 (see Chapter 5) provided some evidence of focus being established prior to an explicit pronominal probe, the self-paced reading study reported by Sanford et al (1996) indicated that complement set anaphors may be processed more slowly than refset anaphors (although this effect was reliable only by materials), suggesting that compset anaphors may be resolved less immediately. Thus, the evidence yielded by previous work is inconclusive. The

experiments reported in the following two chapters represent an attempt to empirically resolve the question.

Chapter Six

The Time Course of Complement Set Focus: A Reading Time Investigation

6.1 Introduction

The introduction to this section related three possibilities for the time course for compset focus. The purpose of this chapter is to test two of these. Specifically, the experiments address the question of whether complement set focus is developed prior or subsequent to an explicit pronominal reference. These experiments make use of an anomaly correcting property which is associated with adversative connectives. The basic rationale which underpins these studies is explained below.

6.1.1 Adversatives and anomaly correction

Connectives, as outlined in Chapter 3, represent a specific class of function word and are held to provide the processor with information concerning the relationship which exists between adjacent pieces of text. Adversative connectives appear to signal a contrasting, or unexpected, relationship between the associated pieces of text. In particular, adversatives would appear to be frequently utilised to signal that a given statement is contrary to some expectation and thereby render acceptable an apparently incongruous statement. This property of adversatives was evident in the continuation types generated in response to the *but* condition of experiment 4, where participants often produced continuations which related the existence of a paradoxical relationship between the quantified statement and the continuation. This general function of adversatives can be illustrated with reference to sentences (1), (2), (3) and (4) below. In these, sentence (1) introduces the quantifier

Tony and relates a property of this, while sentences (2), (3) and (4) make an anaphoric reference to *Tony* and relate a second property.

(1) Tony was worried about the progress of his PhD student.

(2) He was very confident that the student's thesis would be submitted on time.

(3) And so, he was very confident that the student's thesis would be submitted on time.

(4) However, he was very confident that the student's thesis would be submitted on time.

Given the initial statement (1), sentence (2) does not appear to plausibly follow as the properties attached to *Tony* of being *worried about...* and *very confident that...* appear to be inconsistent with each other. When sentence (3) is substituted as the anaphoric reference sentence, the mismatch appears to be amplified as inclusion of the causal connective *and so* provides the processor with the instruction that the second sentence should relate an effect caused by the initial sentence. However, when sentence (4) is the anaphoric reference sentence, the adversative *however* appears to correct the mismatch and render the statement acceptable.

This general pattern of incongruity and acceptability can, in principle, be extended to cases where the initial statement contains a plural set quantifier. Here a central consideration is whether the quantifier is associated with refset or compset focus. Thus, given an initial sentence of (5), for refset focusing *many*, or (6) for compset focusing *not many*, a similar pattern of acceptability and incongruity is maintained for anaphoric reference sentences (7), (8) and (9).

(5) *Many* of the lecturers were **worried** about the progress of their PhD students.

(6) *Not many* of the lecturers were **happy** about the progress of their PhD students.

(7) They were very confident that the students' theses would be submitted on time.

(8) And so, they were very confident that the students' theses would be submitted on time.

(9) However, they were very confident that the students' theses would be submitted on time.

6.1.2 The time course of connective integration

A second issue to be addressed relates to the time course of connective integration. With respect to this, two distinct models have been suggested: the *delayed-integration* model proposed by Millis and Just (1994), and the *incremental* model proposed by Traxler, Bybee and Pickering (1997). The delayed-integration model suggests that when readers encounter a two clause statement separated by a connective, the clauses are processed separately and integration is deferred until after both clauses have been processed. Millis and Just (1994) found evidence to support this model in a series of studies which utilised word-by-word presentation self-paced reading along with a probe recognition task. However, Traxler, Bybee and Pickering (1997) have argued that connectives may be integrated in an incremental way. In general terms, incremental accounts of text comprehension suggest that readers may interpret small fragments of text as they are encountered during reading (e.g. Garrod, Freudenthal & Boyle, 1994; Tyler & Marslen-Wilson, 1977). Traxler et al (1997) investigated the timing of causal connective integration using an eye-movement methodology and found evidence which was consistent with interclause relationships being established incrementally. Given the problematic nature of word-by-word presentation methodologies and the higher relative

sensitivity of the eye-tracking procedure, it would seem reasonable to conclude that connectives are more likely to be integrated incrementally.

6.1.3 The timing of adversative impact as an index of quantifier focus development

Given the information outlined above, two points are salient to the experiments reported in this chapter. First, there is evidence to suggest that connectives are incrementally integrated into the discourse model. If this is the case, adversatives are likely to exert an influence upon comprehension relatively immediately they are encountered. Secondly, for plural set quantifiers, the apparent incongruity of an effect described in an anaphoric reference sentence is dependent upon the focus pattern associated with the quantifier. That is, a statement which is incongruous with a property of a compset will not generally be incongruous with an associated refset property. Thus for the anomaly correcting property of adversatives to facilitate the comprehension of an anomalous statement, and therefore advantage processing of the anomalous sentence, focus must first be established.

It would therefore in principle appear possible to use the timing of the onset of a processing advantage through the inclusion of an adversative as an index of the approximate time at which focus is established. That is, if the adversative is encountered after focus is established, the adversative may confer a processing advantage at any point subsequent to this. However, if the adversative is read before focus is established, any processing advantage should not be apparent until focus is eventually developed. Thus, if an adversative is included at various plausible sites within an anaphoric reference sentence, any relative processing advantage of an early site over a later site would suggest that focus has been established at some point prior to the later adversative being encountered (since advantaged processing is only possible once focus has been established). It is this principle which is utilised by the experiments described in this chapter.

Experiments 5 and 6 test the hypothesis that complement set focus is established subsequent to an initial quantified statement and during the comprehension an anaphoric reference sentence.

6.2 Experiment Five

The Effect of Early versus Late Adversative Placement on Reading Time for an Anomalous Anaphoric Reference Sentence

6.2.1 Introduction

Experiment 5 tests the hypothesis that complement set focus arises only after an explicit pronominal reference is encountered by the reader. Specifically, in it I compare the effect of early (start of sentence) and late (end of sentence) adversative placement on the comprehension of an incongruous anaphoric reference sentence for a monotone increasing set and a monotone decreasing set of quantifiers. Given the assumption of incremental adversative integration, an advantage of early over late placement would suggest that focus is established prior to the late adversative being encountered. An asymmetry in this pattern between the quantifier sets (such that an early advantage holds for the monotone increasing set but not for the monotone decreasing set) would be consistent with focus being established earlier for the monotone increasing set.

6.2.2 Method

Materials and design

Thirty sets of passages (of the type shown in Table 6.1) were used in the experiment, along with 10 filler items which acted as a counterbalance for one of the experimental conditions. Each passage had an informative title and was three sentences long. A full list of these materials is provided in Appendix 2.

Table 6.1 Example experimental and filler materials from Experiment 5

Experimental conditions:

At The Boy-scout Camp

It was late before the boy-scouts had finished setting-up camp.

* **M↑ Q** of the boys felt wide awake after the work.

* **M↓ Q** of the boys felt very tired after the work.

[*And so*][*Nevertheless,*] they slept soundly until morning arrived [, *nevertheless*].

Was the camp set-up by girl-guides?

Filler item:

In Sunday School

The Sunday School class was in full swing.

* **M↑ Q** of the children were unhappy about going to Sunday School.

* **M↓ Q** of the children were happy about going to Sunday School.

And so, they complained to their parents when they returned home.
Were the children attending Sunday School?

Where: **M↑ Q** was one of *all*, *a few* or *many*, and **M↓ Q** was one of *none*, *few* or *not many*.
"*" denotes between groups variants and "[ ]" denotes adversative variants

Experimental items:

The first sentence set the context for the passage and introduced a plural set. The second sentence began with a quantifier and had a verb-phrase which described a situation which was incongruent with the verb-phrase of the following sentence. The quantifier represented the between groups manipulation (*quantified context*) and had six levels (composed of three monotone increasing quantifiers and three monotone decreasing): *all (x), a few (x)* and *many (x)* (the monotone increasing set of quantifiers); *none (x), few (x)* and *not many (x)* (the monotone decreasing and denial set of quantifiers). The situation described by the verb-phrase was modified so as to take account of the focus pattern associated with the quantifier used and maintain semantic similarity across quantifiers. For example, from the passage shown in Table 6.1, "***Many** of the boys felt **wide awake***" and "***Not many** of the boys felt **very tired***" are similarly incongruent with the information that they had subsequently *"slept soundly"* during the night.

The third sentence (*anaphoric reference*) contained a plural pronoun that referred to the preceding quantified noun-phrase and was followed by a verb-phrase (which took the form of an intransitive verb followed by an adverb). The sentence was completed by a temporal prepositional phrase. This sentence appeared in one of three ways, corresponding to the within subjects conditions of: *amplified mismatch; early adversative;* and *late adversative*. In all cases the basic sentence remained the same across conditions (excepting the inclusion of *nevertheless* or *however* in the adversative conditions and *and so* in the amplified mismatch condition).

In the *amplified mismatch* condition an additive/causal connective combination (*And so*) began the sentence to facilitate an amplification of the sentences' incongruence and provide a comparison point for the adversative conditions. In the *early adversative* condition, an adversative connective was placed at the beginning of the sentence and in the *late adversative* condition, the adversative was positioned at the end of the sentence. In half of each adversative condition, the adversative was *however*; in the other half, *nevertheless*.

The experiment, therefore, employed a mixed design (6 x 3). The between-subjects factor (*quantified context* sentence) had six levels: all, none, a few, few, many and not many. The within-subjects factor (*anaphoric reference* sentence) had three levels: amplified mismatch; early adversative; late adversative. The dependent measure was sentence reading time (measured in milliseconds).

The set of 30 experimental materials was divided into 3 blocks (i.e. 10 passages per condition) and rotated, following a Latin Square design, to form 3 experimental files for each quantifier (i.e. a total of 18 files). Five subjects viewed each of the experimental files. Each subject viewed all 30 passages and saw each in only one of the experimental conditions. The order of presentation of the passages was randomised for each subject.

Filler items:

Ten filler passages were included in each of the experimental files as a balance for the amplified mismatch condition. These were of a similar form to the experimental items with the exception that the verb-phrase of sentence two was congruent with that of the target sentence (an example is included in Table 6.1). The causal (*And so*) connective again began the sentence, but in this case signalled a non-anomalous continuation. This was necessary in order to avoid the presence of *And so* acting as a cue for forthcoming anomalous text.

Participants

The participants were ninety students at the University of Glasgow. They were each paid £2.00 for their participation in the experiment. All were native English speakers with normal or corrected to normal eyesight. They were all unaware of the aims of the experiment.

Apparatus

Stimulus materials were presented on a 1024x768 pixel colour monitor controlled by an Apple Macintosh computer running PsyScope (version 1.1.1) experimental design software developed by Cohen, MacWhinney, Flatt and Provost (1993). Participants responded via a button box. This had three buttons, coloured: red, yellow and green (from left to right). The central, yellow, button was used to control the presentation of items and record reading time. The other two buttons recorded subject responses to comprehension questions. Red recorded *NO* responses and green recorded *YES* responses. The button box timer was accurate to one millisecond.

Procedure

Each subject was tested individually. Subjects were seated in front of the computer monitor. Subjects then read a set of instructions presented on the screen and had the opportunity to ask further questions. They were instructed to read at their usual rate and to comprehend what they read as well as possible. Prior to beginning the experiment, subjects were given four practice passages.

Before each trial, a fixation cross appeared at the centre left-hand side of the screen. This was replaced by the title of the passage when the subject pressed the yellow button of the button box. The passages were, likewise, presented sentence-by-sentence at a rate controlled by the subject. There was a comprehension question after each passage. This queried factual information about the first or second sentence in each passage. Half the questions required a YES response, the other half a NO response. There were two opportunities for the subject to have a break during the experiment. Reading time was recorded for sentence three.

6.2.3 Results

Error rate

The overall error rate for the 90 participants was 8.7%. A further breakdown of the error rates associated with each of the conditions is presented in Table 6.2 below.

Table 6.2 Error rates expressed as a percentage of overall responses for the questions associated with each of the experimental conditions

Quantifier	*Early adversative*	*Late adversative*	*Amplified mismatch*	Mean
All (x)	8	8	9.3	8.4
A few (x)	7.3	9.3	7.3	7.9
Many (x)	8.6	6	10.6	8.4
None (x)	10.6	8.6	8	9
Few (x)	10	8	8.6	8.8
Not many (x)	11.3	10	7.3	9.5
Mean	9.3	8.3	8.5	

Sentence reading times

Mean reading times were examined for the third (*anaphoric reference*) sentence. Outliers were initially removed by extracting reading times that fell either below 500ms or above 25s. This accounted for 1.2% of the data. For each participant, outliers were further pruned by removing data which fell beyond 2.5 standard deviations on either side of the mean and replacing these with the cut-off point. This accounted for 2.9% of the data.

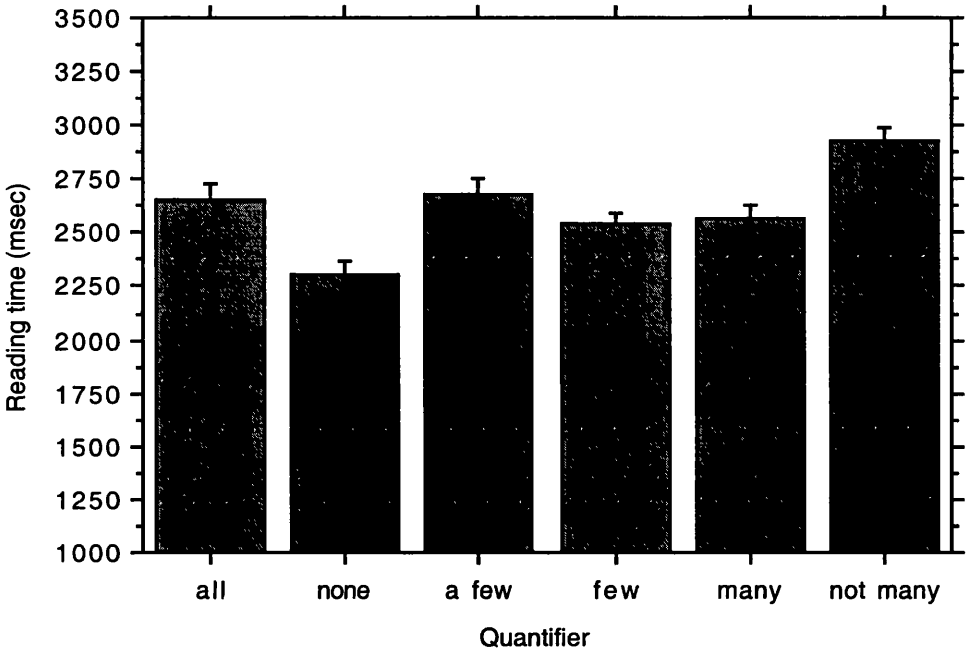
Two 6 x 3 analyses of variance were carried out on the target sentence data. The factors were: the quantifier contained in the previous sentence (*all, none, a few, few, many* and *not many*) and anaphoric reference sentence condition (*early adversative, late adversative* and *amplified mismatch*). The analyses were carried out by subjects (F1) and by materials (F2). The quantifier condition was treated as a between groups variable for the F1 analysis and as a within groups variable for F2.

For quantifier condition, there was no reliable main effect by subjects [$F1(5,84) = 1.385$, $p>0.238$]. There was, however, a reliable effect by materials [$F2(5,29) = 17.766$, $p<0.0001$]. Figure 6.1 relates the overall pattern observed.

A series of simple effects means comparisons were used to investigate the individual quantifier pairs. These indicated that anaphoric references to antecedents quantified by two of the monotone decreasing quantifiers, *none* and *few*, were read reliably faster than references to their monotone increasing counterparts (*none* versus *all*: $F2(1,29) = 26.886$,

p<0.0001; *few* versus *a few*: F2(1,29) = 3.970, p<0.049). For *not many* versus *many*, this pattern was reversed with anaphoric references to sentences quantified by *many* being read more rapidly than references to *not many* [F2(1,29) = 28.109, p<0.0001].

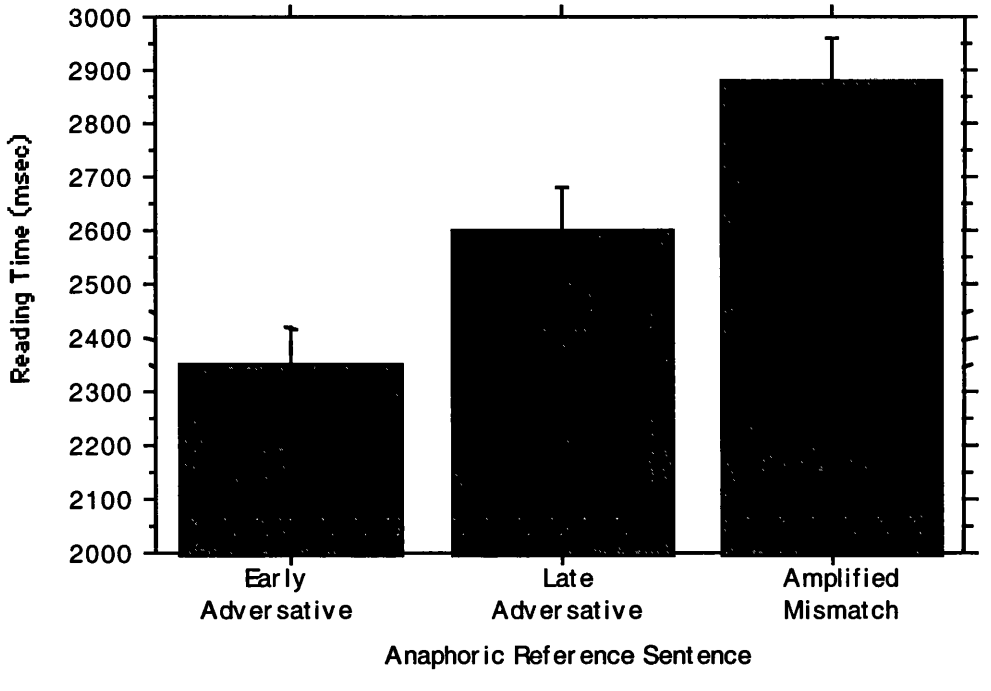
Figure 6.1 Mean reading times in milliseconds (with standard error bars) for the anaphoric reference sentences. The sentences referred to a quantified subset which was introduced in the previous sentence.



Since a major experimental aim was to ascertain whether there exists a difference in the speed of anaphoric resolution between the monotone increasing and decreasing quantifier sets a further simple effects means contrast of these sets was conducted. This analysis indicated that there was no reliable difference between quantifiers divided in this manner [F2(1,29) = 1.173, p>0.280]. The result suggests that there exists no systematic difference in the speed at which compset and refset anaphors are processed. If it can be taken that refset focus is immediately established for monotone increasing quantifiers, as follows from DRT, the result would appear to be consistent with focus having been established prior to the anaphoric reference sentence being encountered for both the monotone increasing and monotone decreasing quantifier sets.

For anaphoric reference sentence conditions, there was a reliable effect both by subjects and by items [F1(2,84) = 43.702, p<0.0001; F2(2,29) = 32.034, p<0.0001]. Figure 6.2 relates the mean reading times which were associated with the anaphoric reference sentence conditions. A series of simple effects means contrasts were carried out to investigate this. These indicated that mean reading times for the early adversative condition were reliably faster than both the late adversative condition (by approximately 256 ms) [F1(1,84) = 20.415, p<0.0001; F2(2,29) = 14.661, p<0.0004] and the amplified mismatch condition (by approximately 531 ms) [F1(1,84) = 87.371, p<0.0001; F2(1,29) = 64.029, p<0.0001]. Additionally, the late adversative condition was found to be reliably faster than the amplified mismatch condition (by approximately 275 ms) [F1(1,84) = 23.319, p<0.0001; F2(1,29) = 17.412, p<0.0001]. This finding indicates that the inclusion of an adversative does appear to advantage the processing of an incongruous sentence (relative to an amplified mismatch). Moreover, the relatively greater advantage of an early adversative over a late adversative would seem to suggest that focus is established prior to the point at which the late adversative is encountered.

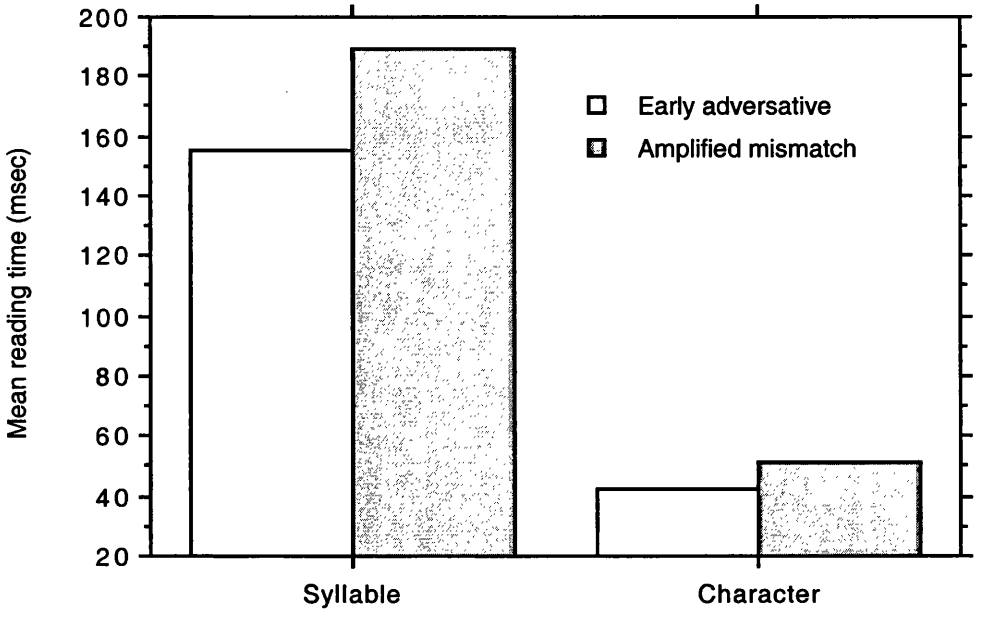
Figure 6.2 Mean reading times in milliseconds (with standard error bars) for the anaphoric reference sentence conditions. The sentences were constructed to relate a property which was incongruous with the focused set and contained either a corrective adversative or causal connective.



There was no reliable interaction between the main effects by subjects or materials [F1(10,84) = 1.182, p>0.306; F2(10,29) = 1.313, p>0.222]. This indicates that there was no asymmetry in the relative advantage/disadvantage associated with the anaphoric reference sentence conditions between the quantifier conditions.

An additional matter to be considered concerns the possible effect of the inclusion of an additional word in the amplified mismatch condition relative to the adversative conditions (i.e. the word pair *and so* in contrast to the single word *however* or *nevertheless*). In order to eliminate this as a potential factor in the observed results, the mean reading times per syllable and per character were calculated for each of the anaphoric reference sentences in the *early adversative* and *amplified mismatch* conditions. The results, presented in Figure 6.3, indicated that higher mean reading times for syllables and characters were associated with the *amplified mismatch* condition. Related measures t-tests indicated that these differences were reliable in both instances (for syllables, t = 6.965, d.f. = 29, $p < 0.005$; for characters, t = 7.007, d.f. = 29, $p < 0.005$). These results suggest that the processing advantage conferred by the inclusion of an adversative is not attributable to addition of an extra word in the *amplified mismatch* condition.

Figure 6.3 Mean reading times per syllable and per character, in milliseconds, for the *early adversative* and *amplified mismatch* anaphoric reference sentence conditions



6.2.4 Discussion

The purpose of Experiment 5 was to test the hypothesis that the establishing of complement set focus is deferred until after an explicit anaphoric reference is encountered. The experiment contrasted the effect of early versus late adversatives on the comprehension of an incongruous anaphoric reference sentence. The property described by the anaphoric reference sentence was refset incongruous for the monotone increasing quantifiers and compset incongruous for the monotone decreasing quantifiers. The central argument is that in order for adversatives to facilitate the comprehension of an incongruous anaphoric reference sentence, the anaphor's textual antecedent must first be identified. Thus, any relative advantage of early over late adversative placement would indicate that focus had been established prior to the late adversative being encountered.

A crucial prediction which arises from the above argument is that if only compset focus is established during comprehension of the anaphoric reference sentence, there should be a differential effect associated with the monotone increasing and monotone decreasing quantifiers. For monotone increasing, there should be an advantage associated with an early adversative, for monotone decreasing there should be less or no advantage.

There were two main findings with respect to the adversative conditions. First, both of these conditions were associated with reliably faster reading times than the amplified mismatch condition. This finding provides further empirical support for the claim that adversatives facilitate the comprehension of anomalous text. Secondly, there was a relative advantage associated with the early adversative condition over the late. This would appear to be consistent with the idea that connectives are incrementally processed, and therefore exert an influence upon comprehension soon after they are encountered. This finding is also consistent with focus having been established prior to the late adversative being encountered. However, it is apparent that there is neither an overall reading time effect associated with monotonicity, nor is there an interaction between the main effects. This would suggest that either both compset and refset focus are established

during the anaphoric reference sentence (but before the late adversative) or, more likely, both have already been established during comprehension of the initial quantified statement.

There is, however, an additional aspect to the data which requires to be addressed. This concerns the observation that, while there was no overall effect of quantifier monotonicity, there was a reliable main effect of quantifier by the materials analysis. Further analysis of this effect indicated that anaphoric references to both *none (x)* and *few (x)* were associated with reliably shorter reading times than were their monotone increasing counterparts (while this pattern was reversed for *many (x)* and *not many (x)*). These findings are surprising in two respects. First, there currently exists no theoretical model which would account for the observed results for *few* and *none*. Indeed, previous research (Sanford et al, 1996) has indicated that anaphoric references to monotone decreasing quantifiers may be read more slowly. Secondly, there is a lack of consistency within the pattern observed (i.e. there is no overall effect of monotonicity). For these reasons, the differences in reading time observed between quantifier conditions appear to be somewhat suspicious and may be spurious.

6.3 Experiment Six

The Effect of Early versus Middle Adversative Placement on Reading Time for an Anomalous Anaphoric Reference Sentence

6.3.1 Introduction

In Experiment 5, no evidence of a systematic asymmetry was found between the monotone increasing and monotone decreasing sets of quantifiers. The most plausible interpretation of this result is that both compset and refset focus is established prior to the anaphoric reference sentence. However, it is also possible that focus for both quantifier types is deferred until some point during comprehension of the anaphoric reference

sentence (but prior to the end of the sentence). The principal purpose of experiment 6 is to test this possibility.

Experiment 6 compares the effect of early and marginally later adversative placement on the comprehension of an incongruous anaphoric reference sentence for the monotone increasing quantifier *many* and the monotone decreasing quantifier *not many*. An advantage for early placement would suggest that focus is established prior to the point at which the later adversative is encountered and thus imply that focus is developed during comprehension of the initial quantified statement. Additionally, the experiment included a single character quantifier condition to act as a baseline comparison point. In this case no ambiguity exists in respect of focused textual entity and this therefore allows for a direct comparison to be made between the effect of adversative position for a case where focus is unambiguously established during the preceding sentence and cases where focus may be established at a later point.

The experiment includes two baseline comparison conditions: one which employs a congruous anaphoric reference sentence and one containing a, non-amplified, incongruous statement. Since the incongruous anaphoric reference sentence does not include the causal connective *and so* in this study, it is possible that the corrective effect of the adversative connective may appear less clear because the reader is not explicitly signalled that the incongruous text should be immediately integrated. If this is the case, it is possible that the disruption caused by the anomalous text may only become clearly evident at a later point. In order to test for this possibility, the experimental items utilised in this study contain an extra, post-target, sentence in order to facilitate the detection of any spill-over effects associated with the conditions.

6.3.2 Method

Materials and design

Thirty-two sets of passages (of the type shown in Table 6.3) were used in the experiment. These were based upon the 30 passages used in Experiment 5 (with two further passages being added). Each passage had an informative title and was four sentences long. A full list of these is provided in Appendix 3.

The first sentence set the context for the passage and introduced a plural set. The second sentence began with a quantifier and had a verb-phrase which described one of two contradictory situations (one congruent with the verb-phrase of the following sentence; the other, incongruent). The quantified noun-phrase represented the between subjects manipulation (*quantified context*) and had three levels: *many, not many* or specified person. The situation described by the verb-phrase was modified so as to take account of the focus pattern associated with the quantifier used and maintain semantic similarity across quantifiers. For example, from the passage shown in Table 6.3, "***Many** of the boys felt **wide awake***" and "***Not many** of the boys felt **very tired***" are similarly incongruent with the information that they had subsequently *"slept soundly"* during the night.

The third sentence (*anaphoric reference*) contained either a singular or a plural pronoun that referred to the preceding quantified noun-phrase and was followed by a verb-phrase (in all but one of the passages this took the form of an intransitive verb followed by an adverb). The sentence was completed by a temporal prepositional phrase. This sentence appeared in one of four ways, corresponding to the within subjects conditions of: *focus match; simple mismatch; early adversative; middle adversative.* In all cases the basic sentence remained the same across conditions (excepting the inclusion of *nevertheless* or *however* in the adversative conditions).

Table 6.3 Example experimental item from Experiment 6

Focus match condition:

At The Boy-scout Camp

It was late before the boy-scouts had finished setting-up camp.

* Many of the boys felt very tired after the work.

* Not many of the boys felt wide awake after the work.

* Simon felt very tired after the work.

They(*He*) slept soundly until morning arrived.

They(*He*) awoke feeling refreshed and ready for the day ahead.

Was the camp set-up by girl-guides?

Simple mismatch condition:

At The Boy-scout Camp

It was late before the boy-scouts had finished setting-up camp.

* Many of the boys felt wide awake after the work.

* Not many of the boys felt very tired after the work.

* Simon felt wide awake after the work.

They(*He*) slept soundly until morning arrived.

They(*He*) awoke feeling refreshed and ready for the day ahead.

Was the camp set-up by girl-guides?

Adversative conditions:

At The Boy-scout Camp

It was late before the boy-scouts had finished setting-up camp.

* Many of the boys felt wide awake after the work.

* Not many of the boys felt very tired after the work.

* Simon felt wide awake after the work.

[*Nevertheless,*] they(*he*) slept soundly[, *nevertheless,*] until morning arrived.

They(*He*) awoke feeling refreshed and ready for the day ahead.

Was the camp set-up by girl-guides?

"*" denotes between subjects variants
"[ ]" denotes adversative variants

In the *focus match* condition, the verb-phrase of the previous (*quantifier*) sentence appeared in a form congruent with the verb-phrase of the anaphoric reference sentence. In

the *simple mismatch*, and both the adversative conditions, the verb-phrase of the quantifier sentence appeared in an incongruent form. In the *early adversative* condition, an adversative connective (*nevertheless* or *however*) was placed before the pronoun. Finally, in the *middle adversative* condition, the adversative was included immediately following the verb-phrase.

The final (*post-target*) sentence of each passage appeared in the same form across conditions. This, again, began with either a singular or plural pronominal anaphoric reference. The sentence conveyed information which was congruent with, but did not explain, the situation described in sentence three.

The experiment, therefore, employed a mixed design (3 x 4). The between groups factor (*quantifier*) had three levels: *many*, *not many* and *individual character*. The within-subjects factor (*anaphoric reference sentence*) had four levels: *focus match*; *simple mismatch*; *early adversative*; and, *middle adversative*. The dependent measure was sentence reading time (measured in milliseconds).

The set of 32 experimental materials was divided into 4 blocks (i.e. 8 passages per condition) and rotated, following a Latin Square design, to form 4 experimental files for each quantified context (i.e. a total of 12 files). Six subjects viewed each of the experimental files. Each subject viewed all 32 passages and saw each in only one of the experimental conditions. The order of presentation of the passages was randomised for each subject. No filler items were presented with the experimental materials.

Participants

The participants were seventy-two students at the University of Glasgow. They were each paid £3.00 for their participation in this and another, unrelated, experiment. All were native English speakers with normal or corrected to normal eyesight. They were all unaware of the aims of the experiment.

Apparatus

Stimulus materials were presented on a 1024x768 pixel colour monitor controlled by an Apple Macintosh computer running PsyScope (version 1.1.1) experimental design software developed by Cohen et al (1993). Participants responded via a button box. This had three buttons, coloured: red, yellow and green (from left to right). The central, yellow, button was used to control the presentation of items and record reading time. The other two buttons recorded subject responses to comprehension questions. Red recorded *NO* responses and green recorded *YES* responses. The button box timer was accurate to one millisecond.

Procedure

Each participant was tested individually. Participants were seated in front of the computer monitor. They then read a set of instructions presented on the screen and had the opportunity to ask further questions. They were instructed to read at their usual rate and to comprehend what they read as well as possible. Prior to beginning the experiment, participants were given four practice passages.

Before each trial, a fixation cross appeared at the centre left-hand side of the screen. This was replaced by the title of the passage when the subject pressed the yellow button of the button box. The passages were, likewise, presented sentence-by-sentence at a rate controlled by the subject. There was a comprehension question after each passage. This queried factual information about the first or second sentence in each passage. Half the questions required a *YES* response, the other half a *NO* response. There were two opportunities for the subject to have a break during the experiment. Reading time was recorded for sentences three and four.

6.3.3 Results

Error rate

The overall error rate for the 72 participants was 8.3%. A further breakdown of the error rates associated with each of the conditions is presented in Table 6.4 below.

Table 6.4 Error rates expressed as a percentage of overall responses for the questions associated with each of the experimental conditions

Quantifier	*Early adversative*	*Middle adversative*	*Simple mismatch*	*Focus match*
Many (x)	8.3	7.3	8.3	6.8
Not many (x)	8.8	8.8	7.8	9.8
Single Character	9.4	6.8	9.4	8.3
Mean	8.8	7.6	8.5	8.3

Sentence reading times

Mean reading times were examined for both the third (*anaphoric reference*) and fourth (*post-target*) sentences. Outliers were initially removed by extracting reading times that fell either below 500ms or above 25s. This accounted for 1.1% of *anaphoric reference* sentence data and 0.7% of the *post-target* sentence data. For each participant, outliers were further pruned by removing data which fell beyond 2.5 standard deviations on either side of the mean and replacing these with the cut-off point. This accounted for 2.4% of the *anaphoric reference* sentence data and 2.1% of the *post-target* sentence data.

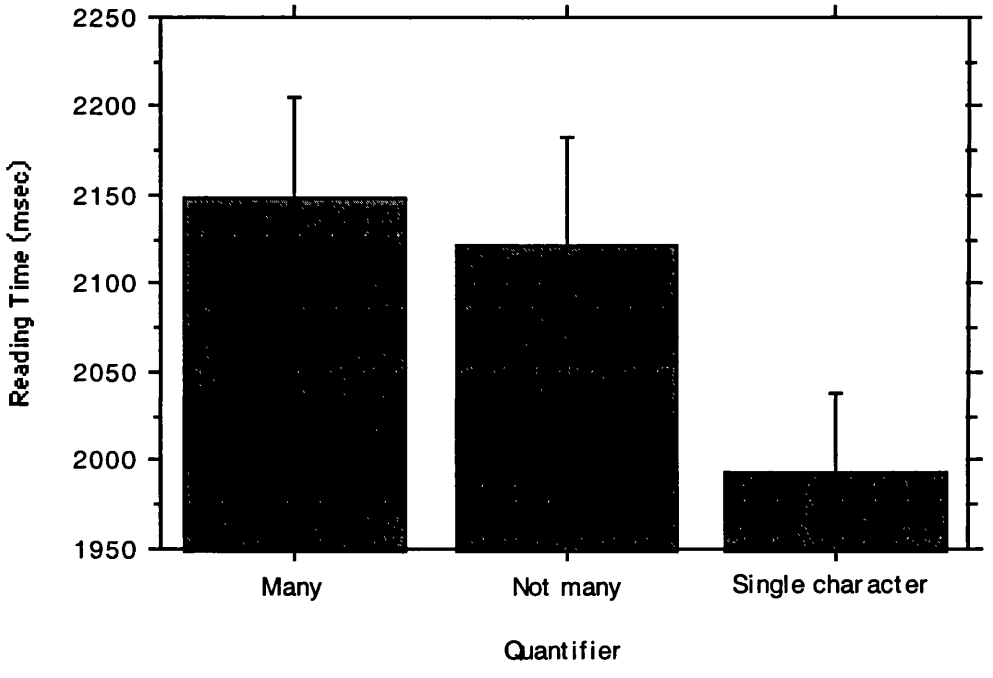
Anaphoric reference sentence analysis

Two 3 x 4 analyses of variance were carried out on the anaphoric reference sentence data. The factors were: the quantifier used in the previous sentence (*many*, *not many* or *single character*) and anaphoric reference sentence condition (*early adversative*, *middle adversative*, *simple mismatch* and *focus match*). The analyses were carried out by

subjects (F1) and by materials (F2). The quantifier condition was treated as a between groups variable for the F1 analysis and as a within groups variable for F2.

For quantifier (see Figure 6.4), while there was no reliable main effect for the by subjects analysis [F1(2,69) = 0.703, p>0.498], the by materials analysis showed a very reliable main effect [F2(2,31) = 11.804, p<0.0001].

Figure 6.4 Mean reading times in milliseconds (with standard error bars) for the anaphoric reference sentences. The sentences referred to a quantified subset which was introduced in the previous sentence.

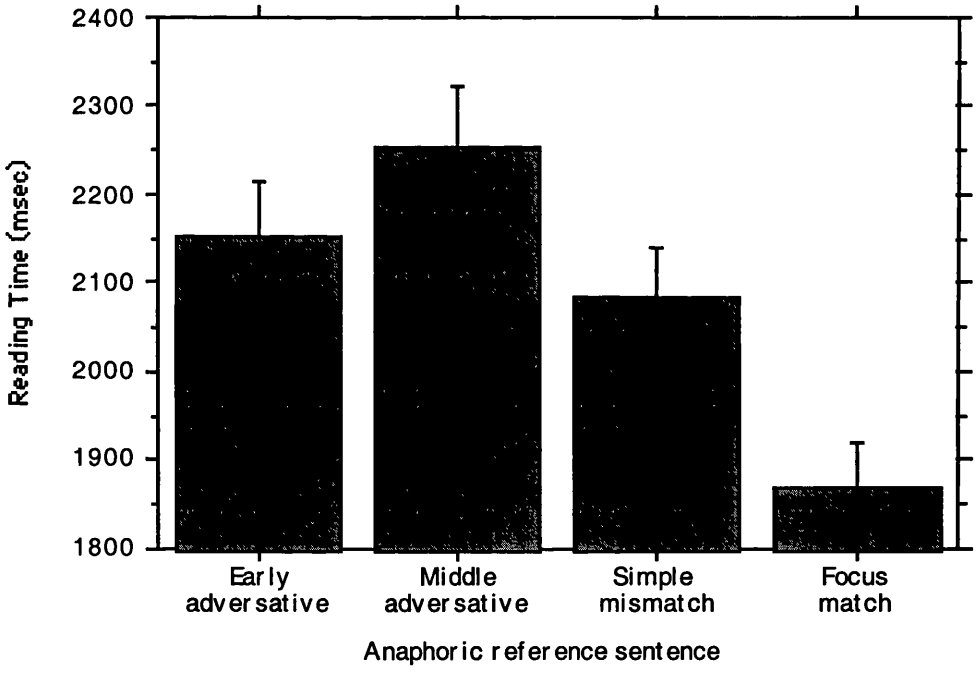


Further investigation of this effect using simple effects means contrast comparisons indicated that the reading times for the sentences containing plural anaphoric references to partitioned sets were reliably slower than those containing a singular pronominal reference to a named individual. The *single character* condition was, on average, approximately 150ms faster than that for *many* [F2(1,31)=19.339, p<0.0001) and approximately 130ms faster than *not many* [F2(1,31)=15.905, p<0.0003]. There was no reliable difference between the sentences containing plural anaphors [F2(1,31)=0.169, p>0.683]. This overall pattern indicates that anaphoric references to plural sets may be processed more slowly than references to single entities. The finding for *many (x)* and *not many (x)*

differs from the pattern observed for these quantifiers in experiment 5 but is, however, consistent with the overall lack of an effect attributable to monotonicity found there. This would seem to further argue in favour of the previous result having been spurious.

There was a reliable main effect for anaphoric reference sentence condition (see Figure 6.5) both by subjects [F1(3,69) = 25.854, p<0.0001] and by materials [F2(3,31) = 15.946, p<0.0001]. A series of means comparisons were conducted to investigate this effect.

Figure 6.5 Mean reading times in milliseconds (with standard error bars) for the anaphoric reference sentence conditions.



For the *early adversative* condition, mean reading times were marginally faster (by approximately 100ms) than the *middle adversative* condition. This difference was reliable by subjects [F1(1,69) = 4.871, p<0.028], but only borderline by materials [F2(1,31) = 3.233, p<0.075]. There was no reliable difference between reading times for the *early adversative* and *simple mismatch* conditions [F1(1,69) = 2.04, F2(1,31) = 1.94; both p's>0.1]. Mean reading times were reliably faster in the *focus match* condition than in: the *early adversative* condition (by about 280ms) [F1(1,69) = 39.404, p< 0.0001; F2(1,31)

= 23.733, p<0.0001], the *middle adversative* condition (by about 380ms) [F1(1,69) = 71.985, p< 0.0001; F2(1,31) = 44.486, p<0.0001], and the *simple mismatch* condition (by about 210ms) [F1(1,69) = 22.346, p< 0.0001; F2(1,31) = 12.105, p<0.0009]. Mean reading time for the *middle adversative* condition was reliably slower (by about 170ms) than for the *simple mismatch* condition [F1(1,69) = 14.117, p< 0.0003; F2(1,31) = 10.18, p<0.002]. To summarise, these data indicate that: sentences in the *middle adversative* condition had the longest mean reading time (2250ms); followed by *early adversative* (2150ms) and *simple mismatch* (2080ms) (which did not reliably differ from each other); and sentences in the *focus match* condition had the shortest mean reading times (1866ms).

These results are interesting in several respects. First, there appears to be an advantage associated with early adversative positioning relative to middle positioning. Again, this pattern is consistent with focus having been established either during the previous sentence or very early during the anaphoric reference sentence. Secondly, the adversative conditions are not associated with faster reading times than the simple mismatch condition. This finding would not appear to be consistent with the view that adversatives facilitate the processing of incongruous statements. However, it is possible that the lack of an obvious benefit is attributable to the longer sentence length associated with the adversative conditions. Thirdly, the *focus match* condition is associated with the fastest reading times which suggests that congruous anaphoric reference sentences are more easily processed than either incongruous or anomaly corrected statements.

There was no interaction between the main effects [F1(6,69) = 0.657, F2(6,31) = 0.417; both p's>0.6]. This indicates that the overall pattern of reading time latencies observed for the anaphoric reference sentence conditions is sensitive to neither the monotonicity of the referent nor whether the referent was a singular or plural set. Given that there is a general advantage of early over middle adversative positioning and that focus is unambiguously established in the previous sentence for the *single character* condition, the lack of an

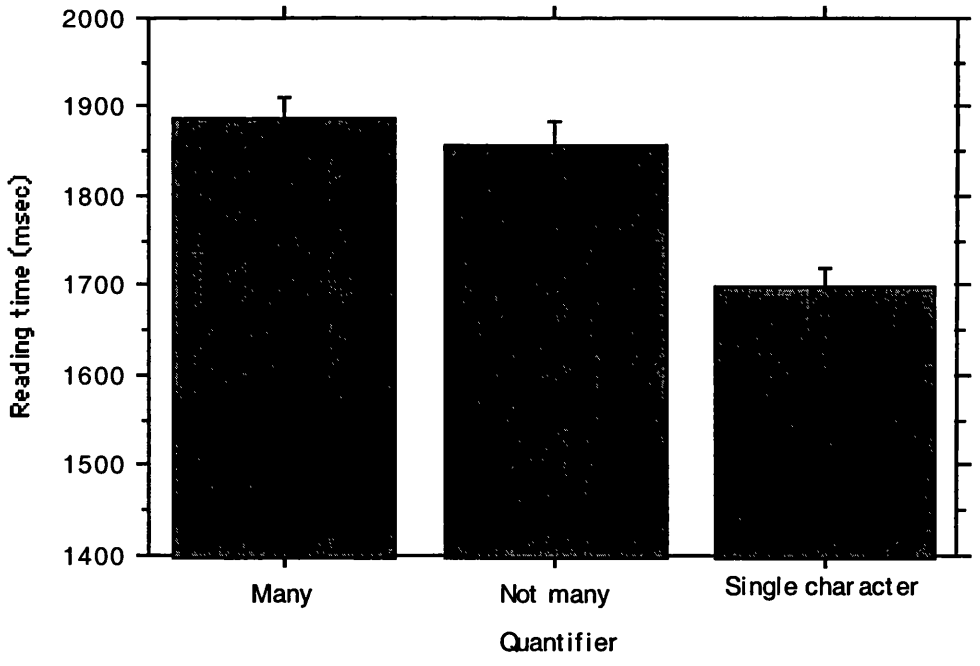
interaction suggests that focus for the plural sets is also established during the preceding sentence.

Post-target sentence analysis

Two 3 x 4 analyses of variance were carried out on the post-target sentence data. The factors were: the quantifier used in sentence two (*many*, *not many* or *single character*) and previous sentence condition (*early adversative*, *middle adversative*, *simple mismatch* and *focus match*). The analyses were carried out by subjects and by materials. The quantifier condition was treated as a between groups variable for the F1 analysis and as a within groups variable for F2.

For quantifier (see Figure 6.6), again there was no reliable main effect for the by subjects analysis [$F1(2,69) = 1.675$]; however, the by materials analysis showed a reliable main effect [$F2(2,31) = 26.324$, $p<0.0001$].

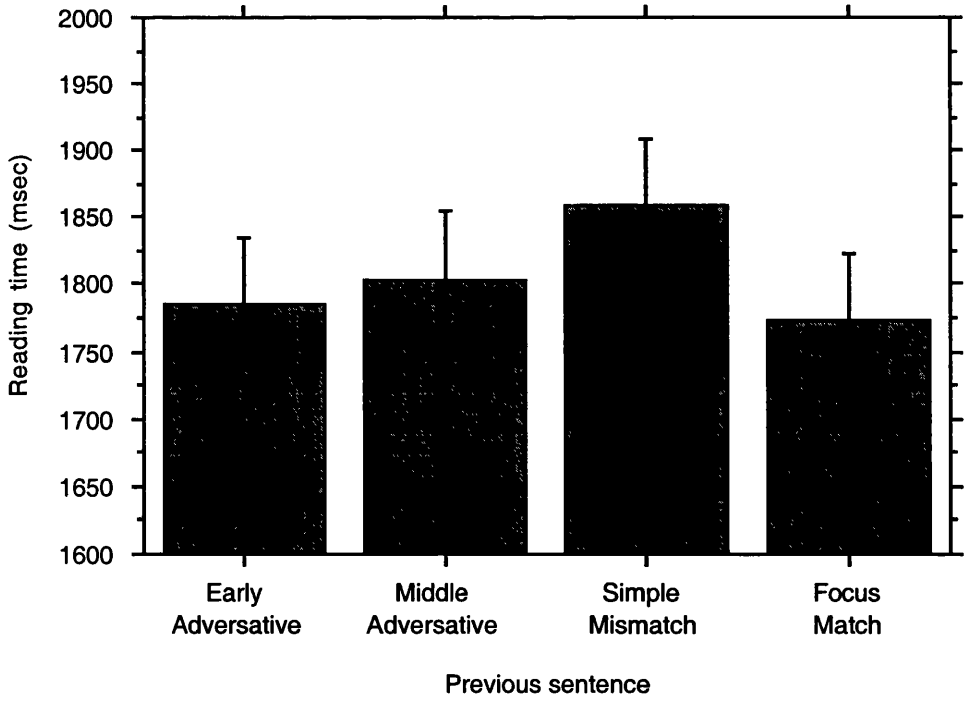
Figure 6.6 Mean reading times in milliseconds (with standard error bars) for the post-target sentences appearing in the three quantifier contexts. The sentence referred to a quantified subset which was introduced in sentence 2.



Further investigation of this effect using simple effects means contrasts again indicated that the reading times for the sentences containing plural anaphoric references to partitioned sets were reliably slower than those containing a singular pronominal reference to a named individual. The quantified context of *single character* was, on average, approximately 200ms faster than that for *many* [F2(1,31) = 45.614, p<0.0001) and approximately 140ms faster than *not many* [F2(1,31) = 32.19, p<0.0001]. There was no reliable difference between the sentences containing plural anaphors [F2(1,31) = 1.167]. This result again indicates that anaphoric references to single named characters are more rapidly processed than references to plural sets.

There was a reliable main effect for previous sentence condition (see Figure 6.7) both, marginally, by subjects [F1(3,69) = 2.597, p<0.054] and by materials [F2(3,31) = 3.063, p<0.032]. A series of planned means comparisons were conducted to investigate this effect.

Figure 6.7 Mean reading times in milliseconds (with standard error bars) for the post-target sentences appearing in the four within-groups conditions. The sentence referred to a quantified subset which was introduced in sentence 2.



There was no reliable difference between mean reading times for sentences preceded by sentences in the *early adversative*, *middle adversative* or *focus match* conditions (*early* vs. *middle* adversatives [F1, F2<0.31]; *early adversative* vs. *focus match* [F1, F2<0.13]; *middle adversative* vs. *focus match* [F1, F2<0.81]). However, sentence reading times for sentences anteceded by the *simple mismatch* condition were reliably longer than those anteceded by the: *early adversative* condition (by about 75ms) [F1(1,69) = 4.935, p<0.028; F2(1,31) = 6.828, p<0.011]; *middle adversative* condition (by about 60ms) for the by materials analysis [F2(1,31) = 4.395, p<0.039], although not by subjects [F1(1,69) = 2.798, p>0.095]; and the *focus match* condition (by about 90ms). [F1(1,69) = 6.615, p<0.011; F2(1,31) = 6.652, p<0.012]. These results indicate that the post-target sentence was equally easily processed when the previous sentence was either congruous or contained an anomaly correcting adversative. However, when the previous sentence was in the *simple mismatch* condition, reading times were reliably slower than in the other conditions. Thus, there appears to be a spill-over effect associated with the *simple mismatch* condition.

There was no interaction between the main effects [F1(6,69) = 0.209, F2(6,31) = 0.645; both p's>0.6].

6.3.4 Discussion

This experiment supports the findings of experiment 5 in three key respects. First, the global time taken to resolve an anaphoric reference to a plural set referent does not appear to be sensitive to quantifier monotonicity (as previously found by Sanford et al, 1996). Secondly, for the adversative conditions, there appears to be a processing advantage associated with early, relative to middle, adversative placement. This finding taken in conjunction with the lack of an interaction between the main effects (and also the lack of an effect attributable to monotonicity) would seem to suggest that complement set focus is established during comprehension of the initial quantified statement. Thirdly, while the adversative conditions are not associated with faster reading times for the anaphoric

reference sentence than the simple mismatch condition, they are associated with faster post-target reading times. This would seem to indicate that while adversatives may not initially benefit the processing of an incongruous statement, they do benefit the integration of subsequent discourse. This evidence provides further support for the anomaly correcting property of adversatives.

6.4 Summary and Conclusions

The experiments reported in this chapter were conducted in order to test the hypothesis that complement set focus is established subsequent to the reading of an initial quantified statement and during the comprehension of an anaphoric reference sentence. In the main experimental conditions of interest, an initial quantified statement was followed by an anaphoric reference sentence which was incongruous with the focus of the previous sentence. Again, in the main conditions of interest, this incongruous sentence contained an anomaly correcting adversative at different sites within the sentence. The central argument was that any relative advantage for an early adversative over a later adversative would suggest that focus had been established prior to the later adversative being encountered by the reader. The main findings were that:

1. The quantifiers studied were associated with a similar pattern of differential focus pattern as has been observed in previous research (Sanford et al, 1996).
2. There was a reading time advantage associated with earlier over later adversative placement. This finding is consistent with an incremental account of adversative processing.
3. There was no evidence to suggest a different time course for the focus effects associated with positive and negative quantifiers.

The general pattern of results obtained in Experiments 5 and 6 would appear to be consistent with focus having been established during comprehension of the quantified

statement and prior to the anaphoric reference sentence. These results are, therefore, equally explicable in terms of both Inference and Set-driven accounts.

Chapter Seven

The Time Course of Complement Set Focus: An Analysis of Eye-movements

Experiment 7[1]

7.1 Introduction

While the evidence from the analysis of the content types generated in the production studies is consistent with the Inference Model, the reading time studies reported in Chapter 6 provided no evidence for an asymmetry in anaphoric resolution between monotone increasing and monotone decreasing quantifiers. The results obtained in these studies were consistent with focus having been established during comprehension of the initial quantified statement as a whole, and, therefore, equally explicable in terms of both Inference and Set-driven accounts. A possible reason for the fact that no systematic effect of monotonicity was detected by the study may lie in the nature of the dependent measure, sentence reading time, used in the study. While this measure is powerful enough to demonstrate general focus mismatch effects (e.g. Sanford et al, 1996), it is not necessarily sufficiently sensitive to differentiate between the models which have been advanced to explain the focus effect. That is, sentence reading time may be too coarse a measure to detect the precise locus of anaphoric resolution.

A more sensitive measure of on-line language processing exists in the monitoring of subjects' eye-movements during reading. The principal assumption underlying this methodology is that there exists a strong connection between the pattern of eye-movements generated in response to a piece of text and the underlying processing of that

[1] This experiment was part of a collaborative effort made by members of the Quantifier Research Group, and was carried out by the author while supported by a British Academy grant awarded to Professor A. J. Sanford. It appeared as Experiment 2 in a paper by Paterson, Sanford, Moxey and Dawydiak (1998).

text (e.g. Liversedge, Paterson & Pickering, 1998; Pollatsek & Rayner, 1990; Rayner, Sereno, Morris, Schmauder & Clifton, 1989). While there has been some debate concerning the degree of link between eye-movements and linguistic processing (e.g. see Rayner et al, 1989; and for a relatively recent exchange, see Vitu, O'Regan, Inhoff & Topolski, 1995 and the subsequent rebuttal by Rayner & Fischer, 1996), there is a considerable volume of evidence which indicates that eye-movements are sensitive to many aspects of linguistic processing. For example, research has shown that eye-movements are sensitive to: the process of anaphoric resolution (e.g. Ehrlich & Rayner, 1983; Garrod, Freudenthal & Boyle, 1994; O'Brien, Raney, Albrecht & Rayner, 1997); the semantic similarity of words (e.g. Carroll & Slowiaczek, 1986); the process of syntactic disambiguation (e.g. Frazier & Rayner, 1982; Rayner, Garrod & Perfetti, 1992) and the lexical ambiguity of words (e.g. Duffy, Morris & Rayner, 1988; Rayner & Frazier, 1989; Sereno, 1995).

A principal strength of this technique lies in the variety of the different individual measures it is possible to derive from the basic record of eye-fixations made during reading. Each of these measures can be individually interpreted as being reflective of a specific aspect of language processing and, therefore, through a combination of the results from these it is possible to build up a detailed picture concerning how a piece of text is processed.

Paterson (1996, experiment three) employed an eye-movement methodology to provide a finer grain measurement of anaphor resolution to a quantified antecedent (*a few* or *few)*. This study made use of the materials used in Sanford et al (1996). In the critical sentence of these materials, the anaphor referred to a property of either the refset or the compset (e.g. *their presence/absence* in the example given in Table 2.7, Chapter 2), thereby allowing the noun-phrase of the critical sentence to match or mismatch with the focused set of the previous quantified sentence.

The results of this study indicated that there was no immediate effect of mismatch for either quantifier. Additionally, for the monotone increasing quantifier (*a few*), the mismatch effect appeared to be stronger both in terms of first pass regressions and in total reading time differences in the anaphoric reference sentence. For the monotone decreasing quantifier (*few*), the mismatch effect was evident only in the total reading time for the final region of the anaphoric reference sentence. This experiment, therefore, provided some evidence for an asymmetry in the processing of anaphors to positive and negative quantifiers.

While the design used in Paterson's experiment was sufficiently powerful to detect difficulty in integrating the anaphoric sentence with an understanding of the preceding text, it is possible that it lacked the power to detect the early mismatch effects predicted by the Set-driven model. In the experiment reported here, manipulations likely to promote the early detection of mismatch effects were used. First, the quantified and anaphoric sentences were conjoined using the causal connective *so*. This connective takes a state of affairs and makes the second state of affairs result from it. Secondly, the potential mismatch was localised to an intransitive verb-phrase rather than to a noun-phrase. Previous work has shown that the main verb region is a good site for detecting the results of pronoun assignments (e.g. Garrod et al., 1994). Taking these changes into account, an example of a mismatch material might be *A few of the men were careful with their winnings, so they gambled recklessly...* Finally, the target sentence was arranged to be the same over all four experimental conditions, so that comparisons could be made over identical regions.

7.2 Method

Materials and Design

Thirty-two sets of passages (an example is given in Table 7.1 and the full list appears in Appendix 4) were used in the experiment. These were based upon the passages used in

Experiments 5 and 6. Each passage had a title and began with a sentence which set the context. The following sentence began with a noun-phrase which was quantified by either *a few* or *few*, and had a verb-phrase which described one of two contradictory situations. In the example in Table 7.1, the verb-phrase describes a situation in which the boy-scouts were either *very tired* or *wide awake.*

In many of the materials, the distinction between the two situations was achieved by using morphologically marked verbs. That is, one verb was marked as the negative of the other. The final sentence of each passage began with a plural pronoun that referred to the preceding quantified noun-phrase, and was followed by a verb-phrase (intransitive verb and adverb) that was congruent with one of the two situations described by the quantified sentence and incongruent with the other. For example, in Table 7.1 the verb-phrase *slept soundly* is congruent with a situation in which the boy-scouts were very tired, but incongruent with a situation in which they were wide awake. The sentence was completed by a temporal prepositional phrase. This had a similar construction across all of the experimental materials. The quantified and anaphoric sentences were conjoined using the causal conjunction *so* in order to mark a causal relation between the two sentences. The passages were double-line spaced.

Table 7.1 An example material from Experiment 7

At The Boy-scout Camp
It was late before the boy-scouts had finished setting-up
camp.# A few/Few of the boys felt very tired/wide awake after the
work, so# they slept soundly# until morning arrived.

Slashes (/) denote alternatives and hashes (#) denote region divisions.

There were two experimental manipulations. The quantified sentence contained either a positive (*a few*) or negative (*few*) quantifier, while the verb-phrase of the quantified

sentence either matched or mismatched with the verb-phrase of the subsequent sentence. This produced a fully-crossed experimental design.

Participants

A total of 36 undergraduate students from the University of Glasgow were paid £5 for their participation in this experiment. All participants were native English speakers and had normal and uncorrected vision.

Procedure

An Stanford Research Institute Dual Purkinje Generation 5.5 eye-tracker was used to monitor the gaze location and movement of subjects' right eye during reading. The eye-tracker has an angular resolution of 1' arc. A Vanilla 386 computer displayed materials on a monitor screen 60 cm from subjects' eyes. The tracker monitored subjects' gaze location every millisecond. The tracker's output was sampled to produce a sequence of eye fixations, recorded as x and y character positions, with their start and finish times.

Before the start of the experiment, subjects read an explanation of the eye-tracking procedure and a set of instructions. They were instructed to read at their normal rate and to read to comprehend the sentences as well as they could. Subjects were then seated at the eye-tracker and placed on a bite-bar and under forehead restraint to minimise head movements. Subjects then completed a calibration procedure.

Before each trial, a fixation cross appeared near the upper left-corner of the screen. Immediately subjects fixated this cross, the computer displayed a target sentence, with the first character of this sentence replacing the fixation cross. This also served as an automatic calibration check, as the computer did not display the text until it detected a stable fixation on the cross. If subjects did not rapidly fixate the cross, the experimenter re-calibrated the eye-tracker. Once subjects had finished reading each sentence, they

pressed a key, and the computer either displayed a comprehension question, or proceeded to the next trial. Comprehension questions followed 25% of the experimental and filler trials. For half of these questions the correct answer was *yes* and for the other half *no*. Subjects responded to the comprehension questions by pressing a *yes* or a *no* response key. There was no feedback on their answers.

The computer displayed each experimental list in a fixed Latin Square order, together with 24 fillers that were materials for another, unrelated, experiment, and an additional 5 filler passages that appeared at the beginning of the experiment and following 3 rest periods.

7.3 Results

Regions and Analysis

The experimental materials were divided into four scoring regions. These are illustrated in Table 7.1. Region 1 contained the title and first sentence. Region 2 contained the quantified sentence and the causal connective *so*. The other two regions divided the anaphoric sentence. Region 3 contained the anaphor (*they*) and the critical intransitive verb-phrase (e.g. *slept soundly*), and region 4 contained the temporal prepositional phrase which completed the anaphoric sentence. The results described are only from the region 3 and 4 analyses, since these were regions following the crucial manipulations. In the one place where an effect in region 2 was found, this is quoted.

Analysis

An automatic procedure pooled short contiguous fixations. Fixations of less that 80ms were incorporated into larger fixations found within one character, and fixations of less than 40ms that were not within three characters of another fixation were deleted. Prior to analysing the eye movement data, those trials where either subjects failed to read the passage properly, or where there had been tracker loss, were removed. More specifically,

those trials were removed in which a zero first pass reading time was recorded for any of the text regions. This accounted for 3.5% of the data.

The eye-movement measures of first pass reading time, total reading time and first pass regressions are reported. *First pass reading time* refers to the total summed duration of eye fixations within a region of text from the time the eye initially enters the region to the time of first exit from that region. That is, it provides a measure of the period of time spent reading a piece of text when it is first encountered. Consequently, first pass reading times have been held to reflect the initial ease of processing associated with a region of text (Liversedge et al, 1998). *Total reading time* for a region of text reflects the summed duration of all fixations made in that region. This measure can therefore be interpreted as reflecting the ultimate ease with which the region is processed and integrated into the textual model. If an effect is found for a region of text for total reading time, but not for first pass reading time, this is generally held to suggest a late effect of the manipulation (Liversedge et al, 1998). *First pass regressions* originating within a region of text are calculated by summing the number of regressive saccades made from the current most rightward fixation within that region. This measure reflects the degree to which left to right movement is disrupted during the first sweep of the eyes through a region of text. The data from each of these measures for each region of text were subjected to two 2 (quantifier type) x 2 (reference type) ANOVAs, one treating subjects as a random variable and the other treating sentences as a random variable.

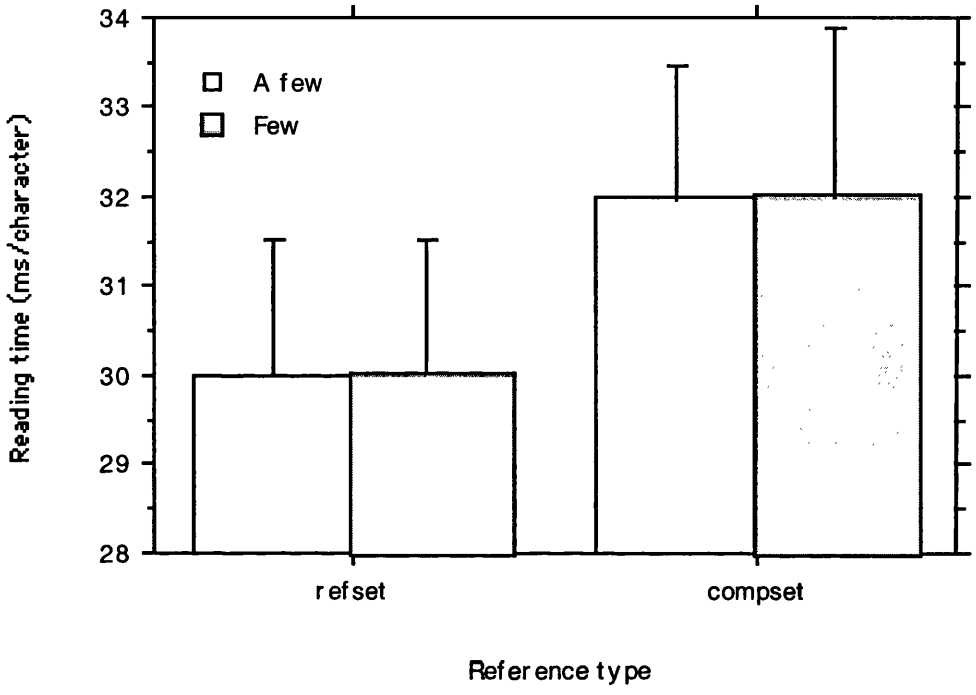
First pass reading time

The mean first pass reading times for regions 3 and 4 are given in Figures 7.1 and 7.2.

At region 3, the critical region, subjects had a significantly longer reading time when the verb-phrase matched the compset as compared to the refset of the preceding sentence ($F1(1, 35) = 6.14$, $p < 0.05$; $F2(1, 31) = 4.38$, $p < 0.05$). There were no other effects

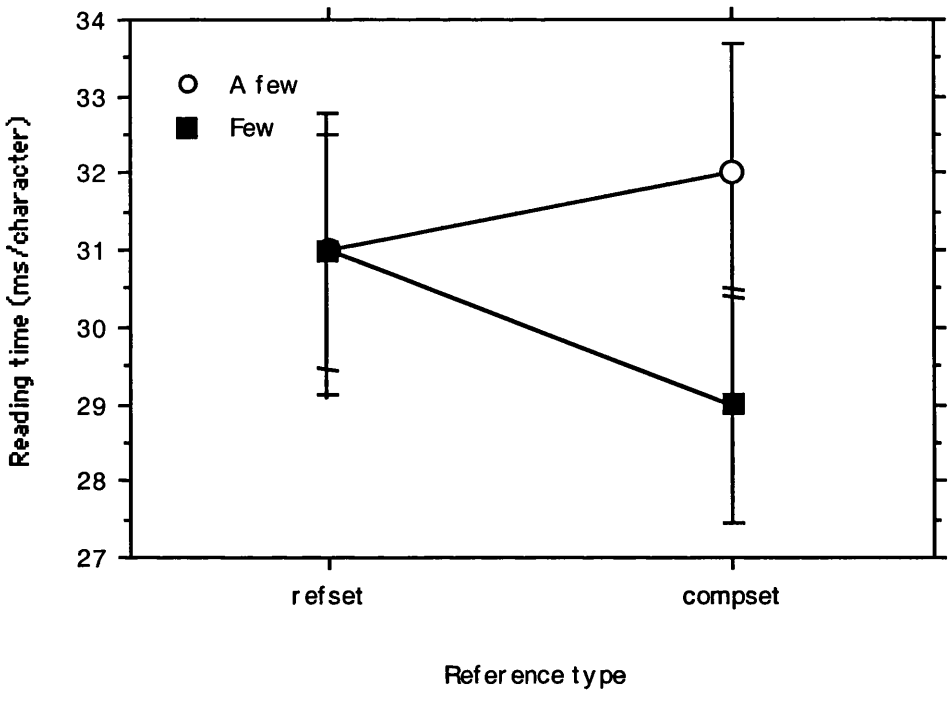
(all Fs < 1). This suggests that subjects had difficulty integrating compset references with the preceding text, irrespective of the polarity of the quantified antecedent.

Figure 7.1 Mean first pass reading times (with standard error bars) for region 3. This region contained the plural anaphor and the critical intransitive verb-phrase.



At region 4, there was a main effect of quantifier type ($F1(1, 35) = 4.79$, $p < 0.05$; $F2(1, 31) = 3.88$, $p < 0.06$), with a longer first pass reading time for this region following quantification by *a few*. There was no main effect of reference type ($F1, F2 < 1$), and no significant interaction of quantifier and reference type ($F1(1, 35) = 2.38$, $p > 0.05$; $F2(1, 31) = 2.04$, $p > 0.05$). This finding is curious as it would seem to suggest a greater initial difficulty in integrating this region of text in the context of quantification by *a few*.

Figure 7.2 Mean first pass reading times (with standard error bars) for region 4. This region contained the temporal prepositional phrase which completed the anaphoric reference sentence.



First pass regressions

The mean number of first pass regressions originating in regions 3 and 4 are given in Figures 7.3 and 7.4.

In region 3, there were no main effects of quantifier type ($F1(1, 35) = 2.17$, $p > 0.05$; $F2 < 1$), reference type ($F1, F2 < 1$), or an interaction of these factors ($F1, F2 < 1$). This would seem to suggest that neither quantifier polarity nor reference type is associated with an initial relative disadvantage in the processing of text in this region.

In region 4, there were no differences in the number of regressions due to either a main effect of quantifier type ($F1, F2 < 1$), or reference type ($F1(1, 35) = 1.6$, $p > 0.05$; $F2(1, 31) = 1.69$, $p > 0.05$). However, there was a significant interaction of these factors ($F1(1, 35) = 7.49$, $p < 0.01$; $F2(1, 31) = 6.84$, $p < 0.05$). There were significantly more regressions when the verb-phrase of the anaphoric sentence matched

Figure 7.3 Mean first pass regressions (with standard error bars) for region 3. This region contained the plural anaphor and the critical intransitive verb-phrase.

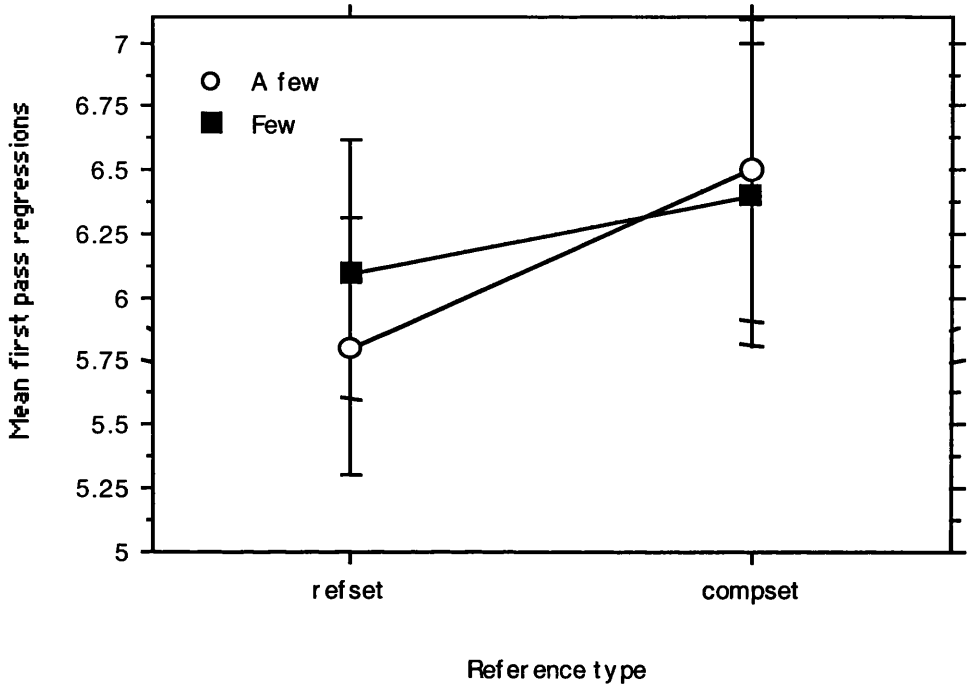
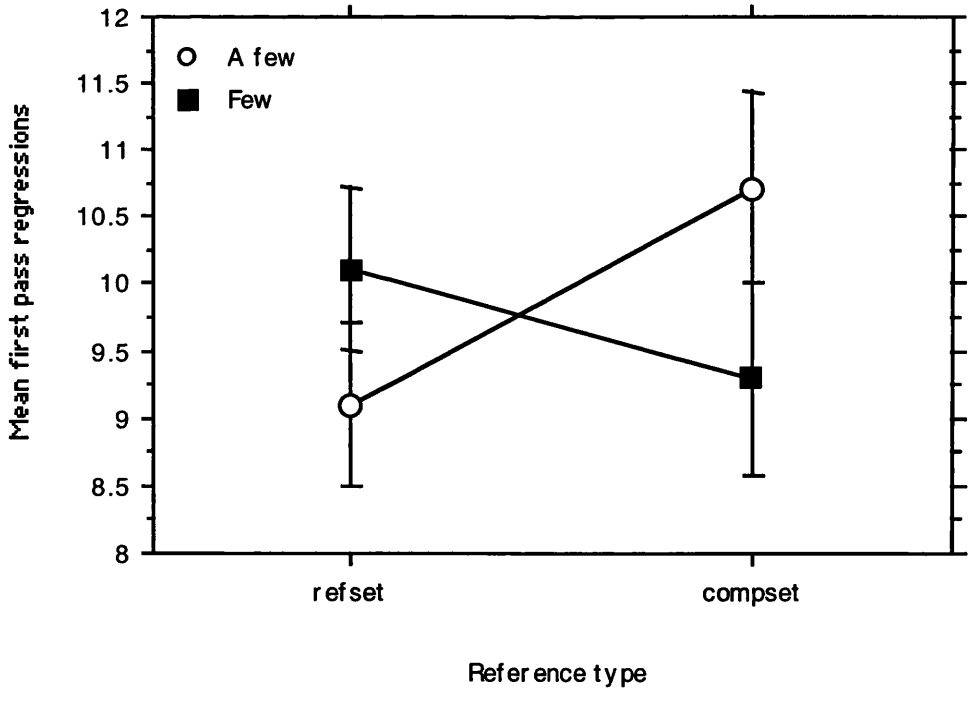


Figure 7.4 Mean first pass regressions (with standard error bars) for region 4. This region contained the temporal prepositional phrase which completed the anaphoric reference sentence.

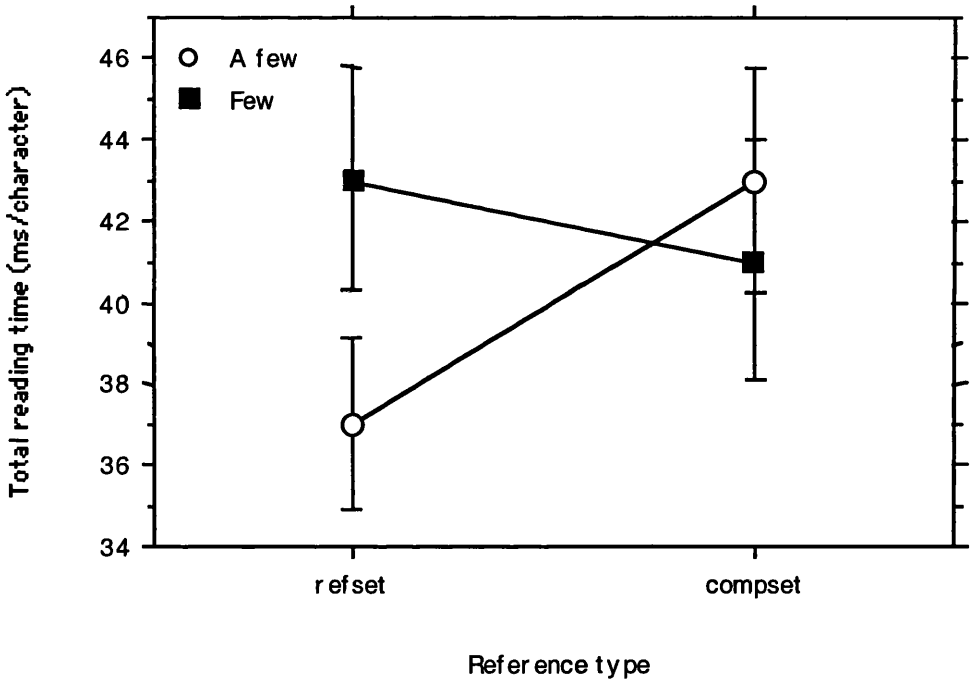


the compset than when it matched the refset following quantification by *a few* ($F1(1, 35) = 6.66$, $p < 0.05$; $F2(1, 31) = 6.54$, $p < 0.05$), but no difference following quantification by *few* ($F1(1, 35) = 1.66$, $p > 0.05$; $F2(1, 31) = 1.10$, $p > 0.05$). This suggests an earlier detection of reference type mismatch following *a few*.

Total reading time

The mean total reading time for regions 3, 4 and 2 are given in Figures 7.5, 7.6 and 7.7.

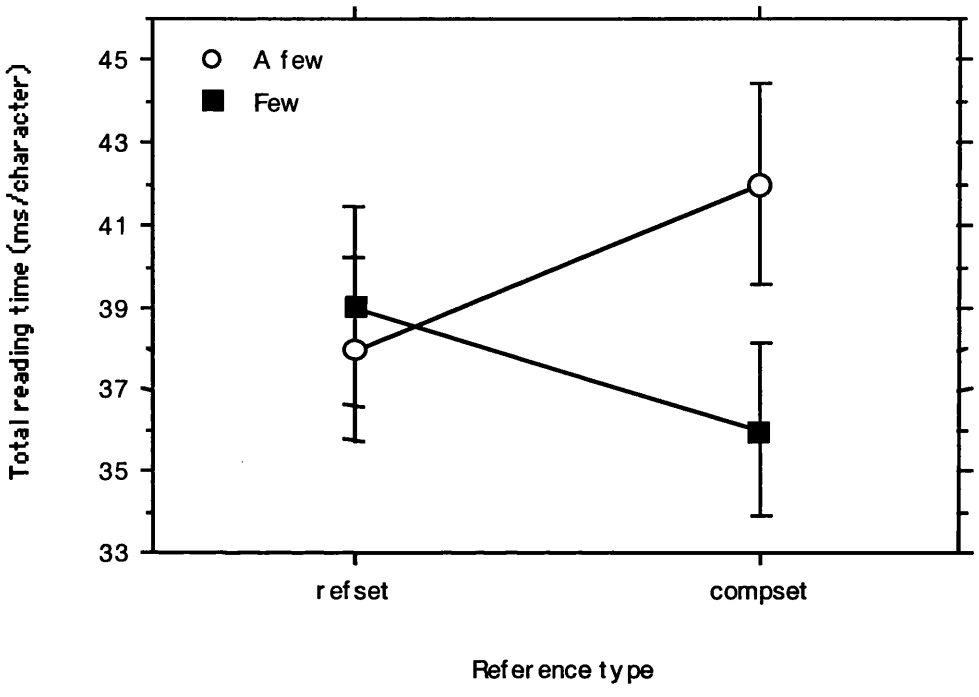
Figure 7.5 Mean total reading time (with standard error bars) for region 3. This region contained the plural anaphor and the critical intransitive verb-phrase.



In region 3, the critical region, there was no main effect of quantifier type ($F1 < 1.8$; $F2 < 1.2$), but there was a main effect of reference type ($F1(1, 35) = 5.46$, $p < 0.05$; $F2(1, 31) = 5.41$, $p < 0.05$). There was also a significant interaction of quantifier and reference type ($F1(1, 35) = 10.91$, $p < 0.01$; $F2(1, 31) = 7.35$, $p < 0.01$). An analysis of the simple effects established that there was a significantly longer reading time for this region when the verb-phrase matched the compset as compared to the refset following quantification by *a few* ($F1(1, 35) = 14.68$, $p < 0.01$; $F2(1, 31) = 10.53$, $p < 0.01$). However, there was

no difference in reading time when the verb-phrase matched either the refset or compset following quantification by *few* (F1, F2 < 1).

Figure 7.6 Mean total reading time (with standard error bars) for region 4. This region contained the temporal prepositional phrase which completed the anaphoric reference sentence.

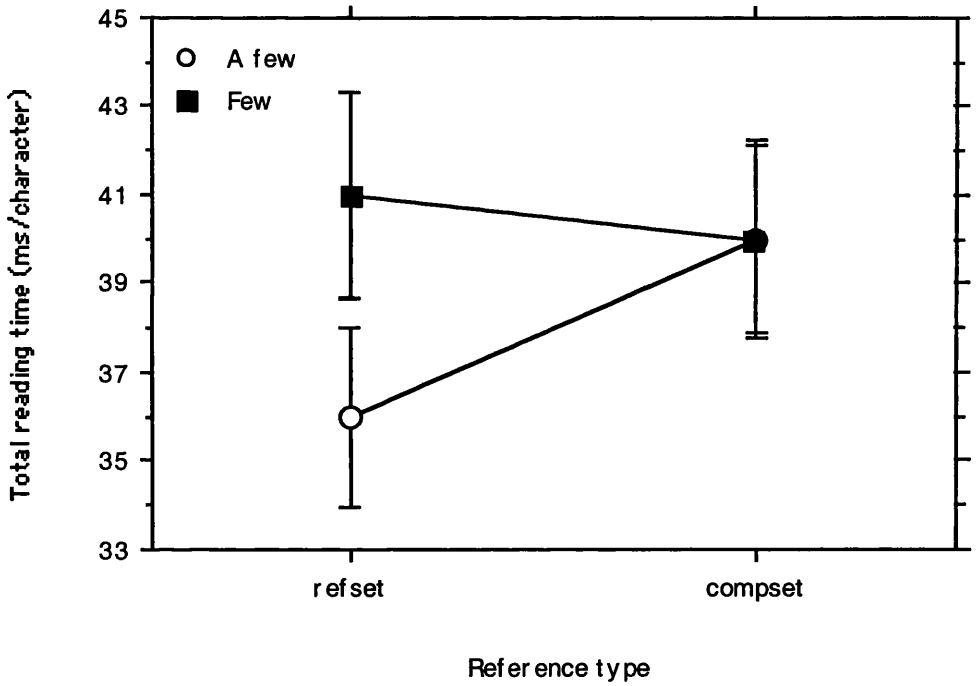


An analysis of the total reading times for region 4 showed a main effect of quantifier type (F1(1, 35) = 8.81, p < 0.01; F2(1, 31) = 7.51, p < 0.01), but no main effect of reference type (F1, F2 < 1). However, there was a significant interaction of these two factors (F1(1, 35) = 10.74, p < 0.01; F2(1, 31) = 12.84, p < 0.01). An analysis of simple effects showed that more time was spent on this region when the preceding verb-phrase matched the compset as compared to the refset following quantification by *a few* (F1(1, 35) = 7.40, p < 0.01; F2(1, 31) = 8.81, p < 0.01). More time was also spent on this region when the preceding verb-phrase matched the refset as compared to the compset following quantification by *few* (F1(1, 35) = 3.67, p < 0.07; F2(1, 31) = 4.41, p < 0.01).

With total time, it is possible for the result of a manipulation to be detected in a region prior to the part of the sentence at which it is introduced, because total time includes fixation times due to second and later passes. In this experiment, such an effect was observed in region 2. There was a main effect of quantifier type (F1(1, 35) = 7.94, p <

0.01; F2(1, 31) = 7.00, $p < 0.05$), but no main effect of reference type (F1(1, 35) = 2.66, $p > 0.05$; F2(1, 31) = 1.05, $p > 0.05$). However, there was a significant interaction of these two factors (F1(1, 35) = 8.02, $p < 0.01$; F2(1, 31) = 5.70, $p < 0.05$). An analysis of the simple effects showed that there was a significantly longer reading time for this region when the verb-phrase of the anaphoric sentence matched the compset as opposed to the refset following quantification by *a few* (F1(1, 35) = 9.29, $p < 0.01$; F2(1, 31) = 7.06, $p < 0.05$). However, there was no difference in total reading time when the verb-phrase matched either the refset or compset following quantification by *few* (F1, F2 < 1).

Figure 7.7 Mean total reading time (with standard error bars) for region 2. This region contained the quantified sentence and the causal connective *so*.



7.4 Discussion

This experiment compared the reading times for regions of text of an anaphoric sentence that contained a plural pronoun followed by an intransitive verb-phrase. The verb-phrase was congruent with either the refset or compset of the preceding quantified sentence.

The results indicated that there was a robust effect on first pass reading time for the region containing the pronoun and intransitive verb-phrase. Specifically, more time was spent initially reading this region of text when the verb-phrase was congruent with the compset. This finding suggests that readers found it difficult to initially process text that referred to the compset of the quantified statement, regardless of the quantifier used. This observation would seem to imply that the focus effects associated with negative quantifiers did not license an initial interpretation of the pronoun as referring to the compset. This finding is inconsistent with the Set-driven model, which claims that focus effects associated with both positive and negative quantifiers drive the process of pronoun resolution. However, it is consistent with a simpler account in which the pronoun is at least initially attached to a referent that is described in the preceding text (i.e. the refset).

The finding that the refset may be the initially activated referent for the pronoun would appear to be interpretable within the framework of the "True" model suggested by Clark (1976). Recall that Clark contends that negative propositions are encoded in terms of their positive counterparts. If this is the case, then it would seem reasonable to suggest that the refset (i.e. the set of entities for whom the predicate is true) may be a necessary part of the encoding negatively quantified sentences such as (1). Here the sentence may be represented as something like: *False (a lot of the men were careful with their winnings).*

(1) Few of the men were careful with their winnings.

If this is the case, then the initial activation of the refset as an antecedent for the pronoun may be attributable to the refset having been activated prior to being negated. That is, if the compset is derived through a process in which the refset is negated, the refset itself would necessarily have to be represented earlier in the process. Thus, if the pronoun is encountered prior to the completed representation of the compset, the pronoun may be initially attached to the refset. The implications of this possibility are discussed in further detail in the concluding chapter.

Consistent with previous findings (Paterson, 1996; Experiment 3), there was evidence in the first pass regression pattern for detection of a focus mismatch in a late region (region 4) when the quantifier was *a few*, but not when it had been *few*. Furthermore, the total reading times showed that in regions 2 and 3 there was a mismatch effect with *a few*, but none with *few*, and a mismatch effect with both in region 4. So while the focus mismatch effect is evident for both quantifiers (consistent with the self-paced reading study reported by Sanford et al, 1996), it occurs later for the negative *few* than for the positive *a few*. That is, for *a few*, the focus mismatch caused a greater initial disruption to sentence processing.

Evidence from this study, then, would seem to suggest the existence of an asymmetry between the resolution of anaphoric reference to positively and negatively quantified antecedents and also between the initial availability of refsets and compsets as referents. This pattern of results would seem to be generally consistent with the Inference Model in that, while there would seem to be an initial processing disadvantage for compset reference, the focus mismatch effect does occur for both quantifiers (although this occurs later for negative antecedents). Comprehension of compset reference would therefore seem to be a less immediate process. This would seem to be consistent with the processor establishing a fit between the virtual question raised by the negative and the content of the anaphoric reference sentence.

Chapter Eight

Summary, Conclusions and General Discussion

8.1 Introduction

The central objective of this thesis has been to empirically evaluate and elaborate the Inference Model of complement set focus. The model suggests that compset reference arises as a result of the types of inferences typically generated in response to quantifiers which signal denial. In more specific terms, expressions of denial are held to signal some form of expectation violation (that is, they deny some existing presupposition). The denial of a presupposition is then thought to lead to a search for reasons which explain why the expected state of affairs did not occur. Compset focus is then generated because reasons of this type will typically relate to some property of those entities for whom the predicate is false (i.e. the compset).

The purpose of this chapter is to provide a summary and general discussion of the central findings of this thesis. I begin by providing an overview of the empirical work reported in the thesis. The chapter then considers the central implications of these findings for the mechanism by which compset reference arises. The chapter concludes by outlining three potential directions for future research.

8.2 A Summary of the Main Experimental Findings

The main experiments reported in this thesis were conducted in order to address three specific aspects of the Inference Model. First, in Experiments 1 and 2 I attempted to empirically evaluate the claim that it is denial, rather than downwards monotonicity, which is the trigger for compset reference. Second, in Experiment 4 I investigated the effect of connective on compset reference. This study was motivated by the idea that

connectives may function to constrain inferential activity and consequently impact upon the incidence of compset reference. The study therefore represents an attempt to directly assess the relationship between the inferences associated with different connectives and the resulting patterns of set focus. Third, Experiments 5, 6 and 7 investigated the time course over which complement set focus is developed. This issue is of considerable importance as the mechanism described by the Inference Model suggests that complement set focus should develop over a more extended period of time than is suggested by Set-Driven accounts.

This section provides a summary of the main experimental findings associated with these areas of enquiry.

8.2.1 Off-line sentence completion studies

Experiment 1 was conducted with two specific objectives. The first was to ascertain whether the patterns of reference and content observed in previous work using explicit pronominal probes would be maintained within the context of a free continuation task. The second aim was to attempt to separate out the effects of two, predominantly co-occurring, properties of compset associated quantifiers: denial and downwards monotonicity. This objective was of considerable importance because the two principal alternative explanations of complement set reference (the Inference Model and Kibble's semantic account) differ with respect to what they consider to be the triggering property of compset focus. The Inference Model states that denial is the trigger, while Kibble's solution claims that it is the property of downwards monotonicity. The principal experimental findings were that:

1/ The continuations generated in this free continuation task repeated the pattern observed in previous (forced reference) work. This finding is of considerable importance in that it demonstrates that compset reference occurs spontaneously and, thus, represents a genuine

phenomenon which is not attributable to the task demands associated with previous forced reference studies.

2/ Pronominal compset reference was largely confined to the monotone decreasing set of quantifiers which formed denials.

3/ Denials were also associated with a higher incidence of RWN content.

In addition to these central findings, the experiment also indicated that there was a high degree of correlation between the apparent strength of denial associated with a given quantifier and the frequency of compset reference. This finding further supports the claim that denial is centrally implicated in the phenomenon and, additionally, provides some initial support for the theory that denial strength may act as a predictor of compset rate and, thus, provide a solution to the problem of variable rates of compset reference.

Experiment 2 represented an attempt to confirm the finding of Experiment 1 that monotone decreasing quantifiers which form denials lead to compset reference and the production of RWN continuations, while monotone decreasing quantifiers which form affirmations do not. As there was no explicit requirement for the participant to refer to a previously introduced set in Experiment 1, this second experiment employed a task which required the participant to refer to a plural antecedent in order to facilitate a clarification of the previous result. The results of Experiment 2 again clearly demonstrated the previously established link between quantifiers which form expressions of denial and compset reference. Additionally, this experiment also provided evidence that complement set reference was equally likely to occur for both proportional and numerical quantifiers. This finding is at odds with Kibble's semantic model which suggests that compset reference should not occur with numerical quantifiers.

An unexpected finding of Experiments 1 and 2 was that the monotone increasing quantifier *no less than 10/10% (x)* was associated with a small, but consistent, incidence

of compset reference. Because this finding was not explicable in terms of either the Inference or Semantic accounts, it was necessary to further investigate the issue. Experiment 3 was conducted in order to address the unexplained finding. It produced evidence which suggested that the effect was attributable to a mismatch between a numerical expectation based on world knowledge and the numerical expectations communicated by the quantifier.

Experiment 4 investigated the impact of connective on focus pattern and content type in order to further examine the relationship between inferential activity and complement set reference. The results indicated that while the connective *but* was found to be associated with reduced incidence of compset reference and RWN production, there was no evidence that the other connectives investigated in the study (*and* and *because*) had any impact upon quantifier associated focus pattern. It was concluded that this pattern of results appeared to be generally consistent with focus having been established prior to the connective being encountered.

8.2.2 On-line studies of anaphoric resolution

Experiments 5 and 6 were self-paced reading time studies which were conducted in order to test the hypothesis that compset focus is established subsequent to the reading of an initial quantified statement and during the comprehension of an anaphoric reference sentence. The experiment employed the anomaly correcting property of adversative connectives as a means of providing an index of the approximate time at which focus is established. The experiments indicated a pattern of differential quantifier focus effects similar to that obtained in previous research (Sanford et al, 1996). There was, however, no evidence to suggest a different time course for the focus effects associated with positive and negative quantifiers. It was concluded that the results appeared to be consistent with focus having been established prior to the reader encountering the plural pronoun for both refset and compset focusing quantifiers.

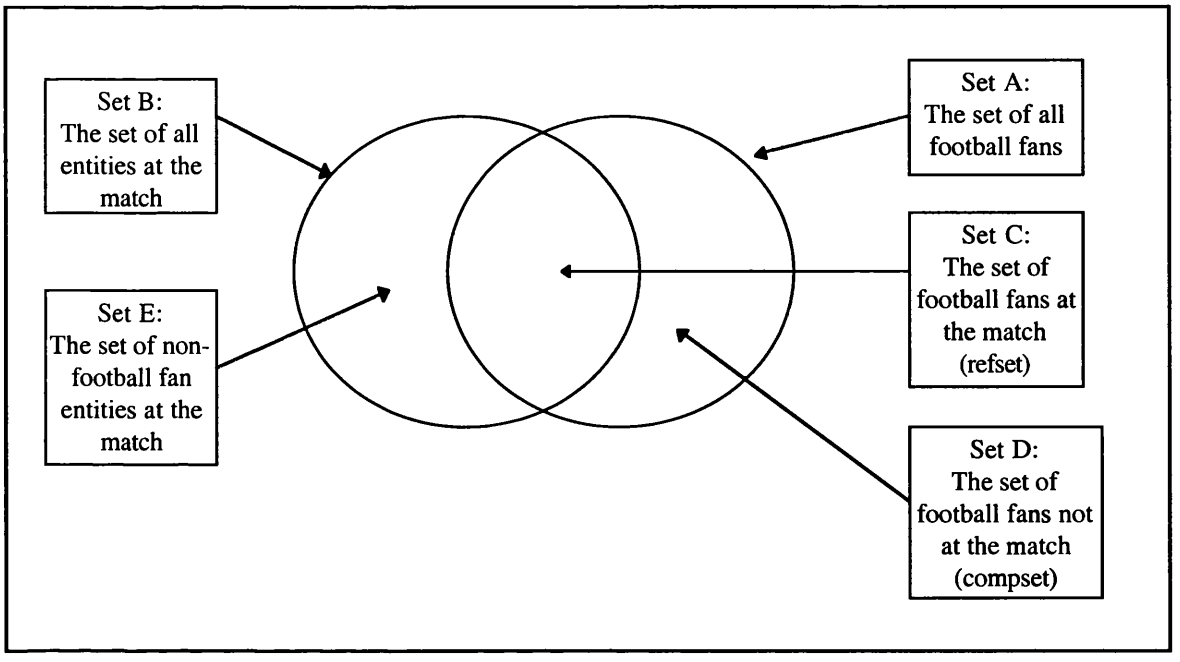
Experiment 7 represented a second attempt to determine the time course over which complement set focus arises. The experiment employed a more sensitive index of on-line linguistic processing (that of eye-movements) in an effort to identify any more subtle differences in the time course of anaphoric resolution for positive and negative quantifiers. The results indicated a focus mismatch effect for both quantifiers. However, unlike Experiments 5 and 6, the results also indicated that the mismatch was detected later for the negative quantifier *few (x)* than for the positive *a few (x)*. This result is consistent with focus having been established earlier for the positive quantifier and is thus generally consistent with the predictions of the Inference Model. An additional finding of this experiment was that the refset may be the initially activated referent for both positive and negative quantifiers. It was argued that this finding may be consistent with the general account of the processing of negativity proposed by Clark (1976).

8.3 General Discussion

The issue of complement set reference can be positioned within the broader context of the representation of quantified statements and the relative accessibility of different sets/subsets within this. Consider a quantified sentence such as (1) below. Here, reasoning would appear to require that five sets are represented (see Figure 8.1). However, in practice, only three of these (i.e. sets "A", "C" and "D") appear to be referred to in the forced reference studies conducted by Moxey and her collaborators. That is, previous research suggests that only these three sets appear to be accessible referents for a plural pronoun. Given this pattern, three main questions were evident at the outset of the research reported in this thesis. First, is compset reference a naturally occurring phenomenon or does it result from the artificial constraints of the forced reference task? Second, if the phenomenon is genuine, what factors impinge upon the relative accessibility of compsets? Third, how do these factors operate in order to bring about complement set reference?

(1) ***Qx*** of the football fans went to the match.

Figure 8.1 The sets represented for the sentence *Qx of the football fans went to the match*



With respect to the initial question posed, Experiment 1 showed that spontaneous complement set reference does occur and thus provided a clear indication of the phenomenon's reality. For the second question, the results of Experiments 1 and 2 strongly suggested that the property of denial (rather than downwards monotonicity) is a central factor in the occurrence of compset reference. The remainder of this section addresses the third question.

8.3.1 Empirical support for the Inference Model

The central objective of this thesis has been to empirically evaluate the Inference Model. In this respect, the research evidence presented in this thesis would appear to be consistent with the Inference account in three respects. First, denial appears to be the primary predictor of compset reference. Second, the central position of denial in the model accommodates a solution to the problem of variable rates of compset reference (since these may be explicable in terms of variable denial strengths being associated with different quantifiers). A third line of evidence which is consistent with this model derives from the finding that focus may be established more immediately for refset focusing

quantifiers than for compset focusing quantifiers. In sum, evidence from this empirical work would appear to be generally consistent with the predictions of the Inference Model and generally inconsistent with the Set-Driven accounts which have been proposed.

8.3.2 An alternative interpretation of the data

While the evidence related above is generally supportive of the Inference account, it is important to consider a second plausible interpretation of the evidence. This alternative interpretation of the data directly follows from Clark's (1976) explanation of the empirical findings reported by Just and Carpenter (1971). Recall that Just and Carpenter used a picture-sentence verification task to investigate the processing of quantified statements (see Chapter 3 for a fuller description of the study). Consistent with the work reviewed and reported in this thesis, Just and Carpenter found evidence to suggest that positively and negatively quantified statements were encoded in different ways: positives being encoded in terms of what corresponds to the refset, negatives being encoded in terms of what corresponds to the compset. A central feature of Clark's (1976) interpretation of this finding is that negatively quantified statements signal a denial of some existing supposition. This denial of an expectation, Clark argues, leads to the negatively quantified statement being encoded in terms of that which it denies. Thus a sentence like (2) from the Just and Carpenter study, might be encoded as *(false(many(dots are red)))*. Here, verification of the truth of such a statement will be based upon an evaluation of the compset (i.e. what it is that is true of *many*).

(2) Not many of the dots are red.

Sanford et al (1996) view such an explanation of Just and Carpenter's findings as being generally consistent with the pattern of variable set focus found in their research. However, they argue that such an account does not entirely explain the pattern of results obtained in their sentence completion studies because refset references are also found to occur in the context of negative quantifiers. Thus, they argue that a more elaborate

account (the initial version of the Inference Model) is necessary to explain the overall pattern of compset and refset reference. However, this argument would appear to be based upon the assumption that negative quantifiers are always interpreted as asserting a denial and thus, in terms of Clark's view, should always be encoded in terms of the compset. If negative quantifiers are not always necessarily interpreted as signalling denial, compset reference would only be expected to result on those occasions where they are interpreted as denial. Such a perspective would appear to provide the basis for an explanation of the incidence of refset reference for negative quantifiers without necessitating the more elaborate approach adopted by the Inference Model.

The alternative perspective outlined above would appear to be consistent with two aspects of empirical evidence presented in this thesis. First, Experiments 1 and 2 demonstrated that the property of denial is centrally implicated in the phenomenon of compset reference. Moreover, Experiment 1 also provided some evidence which indicated that the relative strength of denial may be closely related to the frequency of compset reference for a given quantifier. Secondly, the finding of Experiment 7 that the compset may be derived through a process which involves the negation of the refset is not explicitly predicted by the Inference Model, but is consistent with Clark's explanation of the Just and Carpenter study.

The evidence presented in this thesis would appear to be broadly consistent with both the Inference Model and the alternative account discussed above. Both of these suggest that the property of denial is of central importance to the phenomenon of compset reference. In terms of the Inference Model, denials are thought to lead to a pattern of inferential activity which results in compset reference (here, the incidence of RWN continuations is interpreted as an index of the inferential activity which results in compset reference). In the alternative account, denial is argued to be more directly related to the process by which compsets are derived (here, RWN continuations might reasonably be expected to result from inferential activity which occurs subsequent to focus being established). One

problem for future research, then, may be to develop a way of empirically differentiating between these accounts.

8.4 Unresolved Issues

This section outlines three issues which require further empirical or theoretical clarification.

8.4.1 The relationship between denial strength and compset reference

A potentially important finding of Experiment 1 was the observation that there appears to be a close relationship between the observed frequency of pronominal complement set reference and the strength of denial signalled by a given quantifier. This relationship is of considerable importance as it raises the possibility that the observed variability in the frequencies of compset reference may be explicable in terms of variability in the denial strength associated with different quantifiers. However, the relatively small number of quantifiers used in the study does not allow a strong general claim to be made concerning the relationship. A potential direction for future research, then, may be to systematically gather data relating to the denial strength of an expanded set of quantifiers. Such data would then allow for a clearer indication of the relationship between denial strength and compset reference frequency.

8.4.2 Subjectivity in reference judgements

A common feature of the sentence completion studies, reported both here and elsewhere, is that focus is indexed in terms of judgements concerning the intended referent of a plural pronoun. A potential problem exists for this approach in that these judgements are based upon the subjective interpretations of reference made by judges and participants. The necessarily subjective nature of this type of analysis has enabled claims to be made that the occurrence of compset reference is illusory rather than actual. For example, Corblin

(1996) has suggested that compset reference can be explained as a form of generalisation (see Chapter 2 for a review of Corblin's position). While Moxey and Sanford (2000a) present a series of strong arguments for the reality of complement set reference (see Chapter 2), the subjective nature of the reference judgements represents a potential weakness in the methodology.

One possible method of indexing set focus which removes any reliance on subjective judgements, is suggested by the function of expressions which signal an inclusive relationship. Expressions of this type appear able to function as a signal to the effect that some given entity is a member of a specified set. Thus, in (3) *includes* informs the reader that 3 is a member of the set of prime numbers.

(3) The set of prime numbers *includes* the number 3.

The application of this property of *including* can be illustrated with reference to the quantified text given in (4).

(4) ***Qx*** of the students, *including* Colin, enjoy going to parties.
Did Colin enjoy going to parties?

When a monotone increasing quantifier is substituted for *Qx*, as in (5), the word *including* appears to clearly indicate that *Colin* is a member of the set of students who enjoy parties (i.e. the refset). Thus, in this case, a respondent would be expected to answer "yes" to the question (thereby identifying Colin as a member of the refset).

(5) Almost all of the students, *including* Colin, enjoy going to parties.

Alternatively, when *Qx* is replaced by a monotone decreasing / denial quantifier, as in (6), *including* appears intuitively to identify Colin as a member of the compset. If this is the case, the respondent would be expected to answer "no" to the question.

(6) Almost none of the students, *including* Colin, enjoy going to parties.

In items such as those related above, the pattern of *yes* and *no* responses could be used as an index of which set is the focused entity (*yes* responses would specify refset focus, while *no* responses would indicate compset focus). Thus, future research could, in principle, expand upon this general idea in order to facilitate the investigation of compset reference within a context which removes the need for subjective reference judgements.

8.4.3 Denial and diagnostic tag tests

At one level, the property of denial appears to be clearly defined: denial is an aspect of negativity which signals some manner of expectation violation. Indeed, in a direct test of the quantity expectations signalled by quantifiers, Moxey and Sanford (1993b) showed that the quantity interpreted by participants as having been expected by a communicator was sensitive to the polarity of the quantifier used. Thus, participants interpreted the low ranking (and negative) quantifiers *not many, few, very few* and *only a few* as signalling that a higher amount had been expected relative to the similarly low ranking (but positive) *a few* (see Chapter 3, section 3.6.1 for a more detailed description of the study). In this thesis, however, the empirical demonstration of the relationship between denial and compset reference has relied upon denial being operationally indexed through the use of diagnostic tag tests. The problem is that the linguistic literature contains no clear explanation of why the tag tests are in fact diagnostic of denial (indeed, numerous communications between the Quantifier Research Group and individual linguists have also failed to yield a satisfactory explanation). Thus, the development of a deeper theory of denial would appear to be an important objective for future research.

8.5 Conclusion

A central aspect of this thesis has been to provide a close examination of the relationship between denial and compset reference. In terms of the empirical evidence presented here, it would seem reasonable to conclude that the denial based accounts of complement set reference represent viable explanations of the phenomenon. In contrast, semantic accounts, as exemplified by Kibble's model, do not appear to be consistent with the available evidence.

References

Anscombre, J-C. (1989). Théorie de l'argumentation, topoï et structuation argumenative. *Revue Québécoise de Linguistique,* **18**, 13-65.

Anscombre, J-C. (1991). Dynamique du sens et scalarité. In A. Lempereur (Ed.), *Colloque de Cerisy. L'argumentation.* Liège, Belgium: Mardaga.

Anscombre, J-C., & Ducrot, O. (1986). Argumentation et informativité. In M. Meyer (Ed.), *De la Metaphysique à la rhetorique.* Brussels: Editions de l'Université de Bruxelles.

Barton, S.B., & Sanford, A.J. (1990). The control of attributional patterns by the focusing properties of quantifying expressions. *Journal of Semantics*, **7**, 81-92.

Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy,* **4**, 159-219.

Bass, B.M., Cascio, W.F., & O'Connor, E. J. (1974). Magnitude estimations of frequency and amount. *Journal of Applied Psychology,* **59**, 313-320.

Bem, D.J. (1962). Self perception: An alternative interpretation of cognitive dissonance. *Psychological Review,* **74**, 183-200.

Beyth-Marom, R. (1982). How probable is probable: a numerical translation of verbal probability expressions. *Journal of Forecasting*, **1**, 257-269.

Budescu, D.V., & Wallsten, T.S. (1985). Consistency in the interpretation of probabilistic phrases. *Organizational Behavior and Human Decision Processes*, **36**, 391-405

Carroll, P.J., & Slowiaczek, M.L. (1986). Constraints on semantic priming in reading: A fixation time analysis. *Memory and Cognition*, **14**, 509-522.

Catlin, J., & Micham, D.L. (1975). Semantic representations as procedures for verification. *Journal of Psycholinguistic Research*, **4**, 209-225.

Chafe, W. (1972). Discourse structure and human knowledge. In J. B. Carroll and R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge* (pp.41-69). Washington, DC: Winston.

Chase, C.I. (1969). Often is where you find it. *American Psychologist*, **24**, 1043.

Chien, Y.C. (1994). Structural determinants of quantifier scope: An experimental study of Chinese first language acquisition. In B. Lust & G. Hermon (Eds.), *Syntactic theory and first language acquisition: Cross-linguistic perspectives, Vol. 2: Binding, dependencies, and learnability* (pp. 391-415). Hillsdale, NJ: Erlbaum.

Clark, H. H. (1976). *Semantics and Comprehension.* The Hague: Mouton.

Clifton, C., & Ferreira, F. (1987). Discourse structure and anaphora: Some experimental results. In M. Coltheart (Ed.) *Attention and performance XII: The psychology of reading* (pp. 635-654). Hillsdale, NJ: Erkbaum.

Cohen, J.D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments and Computers,* **25**, 257-271.

Corblin, F. (1997). Quantification et anaphore discursive: la reference aux complementaires. *Langages*, **123**, 51-74.

Devlin, N. (1997). *Pronominal anaphora resolution and the quantified phrase.* Unpublished MA Thesis, UC, Santa Cruz, CA.

Ducrot, O. (1988). Topoï et formes topiques. In *Bulletin d'études de linguistique française,* **2**, 1-14.

Ducrot, O. (1995). Les modificateurs déréalisants. *Journal of Pragmatics,* **24**, 145-166.

Duffy, S.A., Morris, R.K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language,* **27**, 429-446.

Ehrlich, K., & Rayner, K. (1983).Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior*, **22**, 75-87.

Ekberg, P.H.S., & Lopes, L.L. (1979). Syllogistic reasoning: A review. *Goeteborg Psychological Reports*, **9**.

Ekberg, P.H.S., & Lopes, L.L. (1980). Fuzzy quantifiers in syllogistic reasoning. *Goeteborg Psychological Reports*, **10**.

Fodor, J.A., & Garrett, M. (1966). Competence and performance. In: J. Lyons and R.J. Wales (Eds.), *Psycholinguistics Papers*. Edinburgh: Edinburgh University Press.

Frazier, L. & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, **14**, 178-210.

Garrod, S., Freudenthal, D., & Boyle, E. (1994). The role of different types of anaphor in the online resolution of sentences in a discourse. *Journal of Memory and Language,* **33**, 39-68.

Garrod, S., & Sanford, A. J. (1977). Interpreting anaphoric relations: The integration of semantic information during reading. *Journal of Verbal Learning and Verbal Behavior*, **16**, 77-90.

Garrod, S., & Sanford, A. J. (1990). Referential processes in reading: Focusing on roles and individuals. In D.A. Balota, G.B. Flores d'Arcais, and K. Rayner (Eds.), *Comprehension Processes in Reading* (pp 465-485). Hillsdale, NJ: Erlbaum.

Gernsbacher, M.A. (1989). Mechanisms that improve referential access. *Cognition*, **32**, 99-156.

Goodwin, A. R., Thomas, S., & Hartley, J. (1977). Are some parts larger than others? Qualifying Hammerton's quantifiers. *Applied Ergonomics, 8* **(2)**, 9-95.

Gordon, P.C., Grosz, B.J., & Gilliom, L.A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, **17**, 311-347.

Gordon, P.C., & Hendrick, R. (1998). The representation and processing of coreference in discourse. *Cognitive Science*, **22**, 389-424.

Grosz, B. (1977). The representation and use of focus in dialogue understanding. Technical note 15, SRI International Research Center, Stanford, CA.

Halliday, M.A.K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Hammerton, M. (1976). How much is a large part? *Applied Ergonomics,* **7**, 10-12.

Hörmann, H. (1983). The calculating listener or how many are einige, mehrere, and ein paar (some, several and a few)? in R. Bauerle, C. Schwarze, and A. von Stechow (Eds.), *Meaning, Use, and Interpretation of Language*. Berlin: de Gruyter.

Horn, L. R. (1989). *A Natural History of Negation*. Chicago: University of Chicago Press.

Jarvella, R.J., & Lundquist, L. (1994). Scales in the interpretation of words, sentences, and texts. *Journal of Semantics,* **11**, 171-198.

Jarvella, R.J., Lunquist , L., & Hyönä, J. (1995). Text, topos, and Mental Models. *Discourse Processes*, **20**, 1-28.

Johnson, E.M. (1973). Numerical encoding of qualitative expressions of uncertainty. US Army Institute for the Behavioral and Social Sciences, Technical paper 150.

Johnson-Laird, P.N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness*. Cambridge: Cambridge University Press.

Johnson-Laird, P.N., & Byrne, R.M.J. (1989). Only reasoning. *Journal of Memory and Language*, **28**, 313-330.

Johnson-Laird, P.N., & Byrne, R.M.J. (1991). *Deduction*. Hove, UK: Lawrence Erlbaum.

Jones, E.E., & Davis, K.E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (vol. 2, pp. 219-66). New York : Academic Press.

Just, M.A., & Carpenter, P.A. (1971). Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior*, **10**, 244-253.

Kamp, H., & Reyle, U. (1993). *From discourse to logic: Introduction to model-theoretic semantics of Natural language, formal logic and discourse representation theory.* Dordrecht: Kluwer Academic.

Keenan, E.L., & Stavi, J. (1986). A semantic characterization of natural language determiners. *Linguistics and Philosophy*, **9**, 253-326.

Kelley, H.H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on Motivation* (pp. 192-238). Lincoln, NB: University of Nebraska Press.

Kelley, H.H. (1973). The process of causal attribution. *American Psychologist*, 28, 107-28.

Kemtes, K.A., & Kemper, S. (1999). Ageing and resolution of quantifier scope effects. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, **54B**, 350-360.

Kibble, R. (1997). Complement anaphora and monotonicity. In G.J.M. Kruijff, G.V. Morrill. and R. T. Oehrle (Eds.), *Formal Grammar* (pp. 125-136).

Klima, E.S. (1964). Negation in English. In J.A. Fodor and J.J. Katz (Eds.), *The structure of language*. Englewood Cliffs, NJ: Prentice-Hall.

Knott, A. (1996). *A data-driven methodology for motivating a set of coherence relations.* Unpublished Ph.D. Thesis, University of Edinburgh, Scotland, UK.

Kurtzman, H.S., & MacDonald, M.C. (1993). Resolution of quantifier scope ambiguities. *Cognition*, **48**, 243-279.

Ladusaw, W.A. (1979). On the notion of affective in the analysis of negative polarity items. Paper presented to the 1979 Meeting of the Linguistics Society of America, Los Angeles.

Liversedge, S. P., Paterson, K. B., & Pickering, M.J. (1998). Eye movements and measures of reading time. In G. Underwood (Ed.) *Eye Guidance in reading and scene perception.* Oxford: Elsevier Science.

Lundquist, L., & Jarvella, R.J. (1994). Ups and downs in scalar inferences. *Journal of Semantics,* **11**, 33-53.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, **99**, 440-466.

Micham, D.L., Catlin, J., VanDerveer, N.J., & Loveland, K.A. (1980). Lexical and structural cues to quantifier scope relations. *Journal of Psycholinguistic Research*, **9**, 367-377.

Millis, K.K., Golding, J.M., & Barker, G. (1995). Causal connectives increase inference generation. *Discourse Processes*, **20**, 29-49.

Millis, K.K., & Just, M.A. (1994). The influence of connectives on sentence comprehension. *Journal of Memory and Language*, **33**, 128-147.

Moxey, L.M. (1986). *A psychological investigation of the use and interpretation of English Quantifiers.* Unpublished Ph.D. Thesis, University of Glasgow, Scotland, UK.

Moxey, L.M., & Sanford, A.J. (1987). Quantifiers and Focus. *Journal of Semantics,* **5**, 189-206.

Moxey, L.M., & Sanford, A.J. (1993a). *Communicating Quantities.* Hove, UK: Lawrence Erlbaum.

Moxey, L.M., & Sanford, A.J. (1993b). Prior expectations and the interpretation of natural language quantifiers. *European Journal of Cognitive Psychology*, **5**, 73-91.

Moxey, L.M., & Sanford, A.J. (2000a). Focus effects associated with negative quantifiers. In: M.W. Crocker, M. Pickering, and C. Clifton (Eds.), *Architectures and Mechanisms for Language* Processing. Cambridge, UK: Cambridge University Press.

Moxey, L.M., & Sanford, A.J. (2000b). Communicating Quantities: A review of psycholinguistic evidence of how expressions determine perspectives. *Applied Cognitive Psychology*, **14**, 237-255.

Moxey, L.M., Sanford, A.J., & Barton, S.B. (1990). The control of attentional focus by quantifiers. In: K.J. Gilhooly, M. Keane, R.H. Logie, and G. Erdos, *Lines of Thinking,* Vol 1, Chichester: Wiley.

Murray, J.D. (1997). Connectives and narrative text: The role of continuity. *Memory and Cognition*, **25**, 227-236.

Newstead, S.E. (1995). Gricean implicatures and syllogistic reasoning. *Journal of Memory and Language,* **34**, 644-664.

Newstead, S.E., & Collis, J.M. (1987). Context and the interpretation of quantifiers of frequency. *Ergonomics*, **30**, 1447-1462.

Newstead, S.E., & Coventry, K.R. (2000). The role of context and functionality in the interpretation of quantifiers. *European Journal of Cognitive Psychology*, **12**, 243-259.

Newstead, S.E., Pollard, P., & Reizbos, D. (1987). The effect of set size on the interpretation of quantifiers used in rating scales. *Applied Ergonomics,* **18**, 178-182.

O'Brien, E.J., Raney, G.E., Albrecht, J.E., & Rayner, K. (1997). Processes involved in the resolution of explicit anaphors. *Discourse Processes*, **23**, 1-24.

Parducci, A. (1968). Often is often. *American Psychologist,* **23**, 828.

Paterson, K.B. (1996). *The time course of processing of natural language quantification.* Unpublished Ph.D. Thesis, University of Glasgow, Scotland, UK.

Paterson, K. B., Sanford, A. J., Moxey, L. M. & Dawydiak, E. J. (1998). Quantifier polarity and referential focus during reading. *Journal of Memory and Language*, **39**, 290-306.

Pepper, S., & Prytulak, L. S. (1974). Sometimes Frequently means Seldom: Context Effects in the Interpretations of Quantitative Expressions, *Journal of Research in Personality,* **8**, 95-101.

Pohl, N.F. (1981). Scale considerations in using vague quantifiers. *Journal of Experimental Education,* **49**, 235-240.

Pollatsek, A., & Rayner, K. (1990). Eye movements and lexical access in reading. In D.A. Balota, G.B. Flores d'Arcais, and K. Rayner (Eds.), *Comprehension Processes in Reading* (pp 143-164). Hillsdale, NJ: Erlbaum.

Potts, G.R., Keenan, J.M., & Golding, J.M. (1988). Assessing the occurrence of elaborative inferences: Lexical decision versus naming. *Journal of Memory and Language*, **27**, 399-415.

Poulton, E. C. (1973). Unwanted range effects from using within-subject experimental designs. *Psychological Bulletin*, **80**, 113-121.

Poulton, E. C. (1989). *Bias in quantifying judgements.* Hove, UK: Erlbaum.

Rapoport, A., Wallsten, T. S., & Cox, J. A. (1987). Direct and indirect scaling of membership functions of probability phrases. *Mathematical modelling,* **9**, 397-417

Rayner, K., & Fischer, M.H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception and Psychophysics*, **58**, 734-747.

Rayner, K., & Frazier, L. (1989). Selection mechanisms in reading lexically ambiguous words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 779-790.

Rayner, K., Garrod, S.C., & Perfetti, C.A. (1992). Discourse influences during parsing are delayed. *Cognition*, **45**, 109-139.

Rayner, K., Sereno, S.C., Morris, R.K., Schmauder, A.R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes [Special issue]. *Language and Cognition Processes*, 4, 21-49.

Revlis, R., Lipkin, S.G., & Hayes, J.R. (1971). The importance of universal quantifiers in a hypothetical reasoning task. *Journal of Verbal Learning and Verbal Behavior*, **10**, 86-91.

Rosenthal, D.J. (1980). Logical quantifiers in affirmative and negative contexts. *Journal of General Psychology*, **102**, 99-106.

Sanford, A.J., & Garrod, S.C. (1981). *Understanding Written Language*.. Chichester, UK: John Wiley and Sons.

Sanford, A. J., Moar, K., & Garrod, S. (1988). Proper names as controllers of discourse. *Language and Speech*, **31**, 43-56.

Sanford, A. J., Moxey, L. M., & Paterson, K. B. (1996). Attentional focusing with quantifiers in production and comprehension. *Memory and Cognition*, **24**, 144-155.

Schachter, S. (1964). The interaction of cognitive and physiological determinants of emotional state. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (vol. 1, pp. 49-80). New York: Academic Press.

Schriesheim, C., & Schriesheim, J. (1974). Development and empirical verification of new response categories to increase the validity of multiple response alternative questionnaires. *Educational and Psychological measurement,* **34**, 877-884.

Schriesheim, C., & Schriesheim, J. (1978). The invariance of anchor points obtained by magnitude estimation and pair-comparison treatment of complete ranks scaling procedures: An empirical comparison and implications for validity of measurement. *Educational and Psychological Measurement,* **38**, 977 - 983.

Sereno, S.C. (1995). Resolution of lexical ambiguity: Evidence from an eye movement priming paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 582-595

Stewart, A.J., Pickering, M.J., & Sanford, A.J. (2000). The time course of the influence of implicit causality information: Focusing versus integration accounts. *Journal of Memory and Language,* **42**, 423-443.

Taylor, I., & Taylor, M.M. (1990). *Psycholinguistics: Learning and using language.* Eaglewood Cliffs, NJ: Prentice-Hall.

Trabasso, T., Rollins, H., & Shaughnessy, E. (1971). Storage and verification in processing concepts. *Cognitive Psychology*, **2**, 239-289.

Traxler, M.J., Bybee, M.D., & Pickering, M.J. (1997). Influence of connectives on language comprehension: Eye-tracking evidence for incremental interpretation. *Quarterly Journal of Experimental Psychology*, **50A,** 481-497.

Tversky, A., & Kahneman, D. (1974). Judgements under uncertainty: heuristics and biases. *Science*, **185**, 1124-1131.

Tyler, L.K., & Marslen-Wilson, W.D. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, **16**, 683-692.

Van den Berg, M. (1996). Dynamic generalised quantifiers. In J. van der Does and J. Van Eijck (Eds.), *Quantifiers, Logic and Language*. Stanford: CSLI.

Van der Wouden, T. (1997). *Negative Contexts*. London: Routeledge.

Vitu, F., O'Regan, J.K., Inhoff, A.W., & Topolski, R. (1995). Mindless reading: Eye movement characteristics are similar in scanning letter strings and reading text. *Perception and Psychophysics*, **57**, 352-364.

Vonk, W., Hustinx, L.G.M.M., & Simons, W.H.G. (1992). The use of referential expressions in structuring discourse. *Language and Cognitive Processes*, **7**, 301-333.

Wallsten, T. S., Budescu, D. V., Rapoport, A., Zwick, R., & Forsyth, B. (1986). Measuring the Vague Meanings of Probability Terms, *Journal of Experimental Psychology: General,* **115**, 348-365.

Wallsten, T. S., Fillenbaum, S., & Cox J. A. (1986). Base rate effects on the Interpretations of Probability and Frequency Expressions, *Journal of Memory and Language,* **25**, 571-587.

Wason, P.C. (1961). Response to affirmative and negative binary statements. *British Journal of Psychology*, **52**, 133-142.

Wason, P.C. (1965). The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior*, **4**, 7-11.

Wason, P.C. (1971). In real life negatives are false. *Communication and Cognition*, **4**, 239-253.

Wason, P.C., & Jones, S. (1963). Negation: Denotation and connotation. *British Journal of Psychology*, **54**, 299-307.

Wason, P.C., & Reich, S.S. (1979). A verbal illusion. *Quarterly Journal of Experimental Psychology*, **52**, 133-142.

Zwarts, F. (1996). Facets of negation. In J. van der Does and J. van Eijk (Eds.), *Quantifiers, Logic and Language*. Stanford: CSLI.

Appendix 1:

Experimental materials for Experiments 1 to 4

The following list relates the experimental materials used in Experiments 1 to 4. The full set of 30 materials was utilised in Experiments 1 and 4, while Experiments 2 and 3 used only the first 20 in the list. In Experiment 1, *Qx* was one of: *A few, Nearly all, At least 10%, No less than 10%, Few, Not quite all, At most 10%* or *At most 90%* (in the main experimental conditions); *more than 10%* or *less than 10%* (in the auxiliary conditions). In this study, *[Connective]* simply consisted of a space for the sentence completion. In Experiment 2, *Qx* was one of: *At least 10/10%, No less than 10/10%, At most 10/10%* or *No more than 10/10%* (the slashes separate the numerical and proportional forms) and *[Connective]* was replaced by ". They". In Experiment 3, *Qx* was one of: *No less than 80* or *No less than 80%* and *[Connective]* was replaced by ". They". In Experiment 4, *Qx* was one of: *Few, Only a few* or *Not quite all* and *[Connective]* was one of: . *They, and they, but they* or *because they.*

Qx of the MPs went to the meeting ***[Connective]***

Qx of the football fans went to the game ***[Connective]***

Qx of the students produced good essays ***[Connective]***

Qx of the soldiers were overweight ***[Connective]***

Qx of the children enjoyed the fair ***[Connective]***

Qx of the hospital patients volunteered for the test ***[Connective]***

Qx of the job applicants did well in the interviews ***[Connective]***

Qx of the accountants took up the retraining offer ***[Connective]***

Qx of the patients liked the new hospital food ***[Connective]***

Qx of the miners believed the new offer on job safety ***[Connective]***

Qx of the athletes trained for two hours each day ***[Connective]***

Qx of the swimmers had good backstroke technique ***[Connective]***

Qx of the schoolboys finished the marathon ***[Connective]***

Qx of the audience clapped when the play finished ***[Connective]***

Qx of the congregation joined in the singing ***[Connective]***

Qx of the parents allowed their children to go to the club ***[Connective]***

Qx of the girl guides missed the swimming practice ***[Connective]***

Qx of the boys on the estate had taken drugs ***[Connective]***

Qx of the couples at the dance could do the new waltz ***[Connective]***

Qx of the teachers supported the candidate for the union ***[Connective]***

Qx of the Wembley crowed cheered the unusual service ***[Connective]***

Qx of the of the rock-show audience admired the lead guitarist ***[Connective]***

Qx of the purchasers enjoyed the new book ***[Connective]***

Qx of the new doctors had been trained in communication skills ***[Connective]***

Qx of the club members came to the AGM ***[Connective]***

Qx of the Indians on the reservation had a healthy diet ***[Connective]***

Qx of the of the holiday-makers lay on the beach ***[Connective]***

Qx of the people on the train slept sometime during the journey ***[Connective]***

Qx of the general airport staff supported the traffic control strike ***[Connective]***

Qx of the parents liked the headmaster's ideas ***[Connective]***

Appendix 2:

Experimental and filler materials for Experiment 5

Experimental items:

The following list relates the 30 experimental materials used in Experiment 5. In these, "*" denotes the between groups variants. In these, the determiners *None*, *Few* and *Not many* were substituted for **M↓ Q** and *All*, *A few* and *Many* for **M↑ Q**. The within-groups variants are denoted by "[ ]".

At The Boy-scout Camp
It was late before the boy-scouts had finished setting-up camp.
* **M↓ Q** of the boys felt very tired after the work.
* **M↑ Q** of the boys felt wide awake after the work.
[And so][Nevertheless], they slept soundly until morning arrived[, nevertheless].
Was the camp set-up by girl-guides?

In A Retirement Home
All of the residents were in the TV room during Brookside.
* **M↓ Q** of the residents were excited by the plot.
* **M↑ Q** of the residents were bored by the plot.
[And so][Nevertheless], they watched intently when the programme began[, nevertheless].
Were the residents watching Brookside?

Caught Shop-Lifting
A group of school-boys were suspected of stealing from the local shop.
* **M↓ Q** of the boys appeared guilty of the crime.
* **M↑ Q** of the boys appeared innocent of the crime.
[And so][Nevertheless], they confessed immediately when the police arrived[, nevertheless].
Were the boys suspected of vandalism?

At The Theatre
The local theatre company put on a controversial work.
* **M↓ Q** of the audience felt pleased by the performance.
* **M↑ Q** of the audience felt disgusted by the performance.
[And so][Nevertheless], they applauded appreciatively when the play finished[, nevertheless].
Did the theatre company put on a controversial work?

At A Conference
The conference delegates used the university cafeteria for lunch.
* **M↓ Q** of the delegates appeared disappointed by the food.
* **M↑ Q** of the delegates appeared impressed by the food.
[And so][Nevertheless], they complained bitterly when a manager arrived[, nevertheless].
Did the conference delegates go to the pub for lunch?

Keeping Fit
The security guards were almost all unfit and overweight.
* **M↓ Q** of the men were enthusiastic about improving their health.
* **M↑ Q** of the men were apathetic about improving their health.
[And so][However], they exercised vigorously before their shift began[, however].
Were the security guards unfit and overweight?

In A Hospital
Some medical students had the chance to participate in a drugs trial.
* **M↓ Q** of the students appeared reluctant to participate.
* **M↑ Q** of the students appeared keen to participate.
[And so][However], they refused pointedly when the medical authorities asked[, however].
Were the students studying law?

Going To A Party
Some of the school-boys asked their parents for beer to take to the party.
* **M↓ Q** of the parents felt happy about underage drinking.
* **M↑ Q** of the parents felt upset about underage drinking.
[And so][However], they agreed willingly when the boys asked[, however].
Did the schoolboys ask their parents for beer?

In The Classroom
The teacher lectured the entire class about behaving properly.
* **M↓ Q** of the children were scolded by the teacher.
* **M↑ Q** of the children were praised by the teacher.
[And so][However], they wept tearfully until their parents arrived[, however].
Was the class lectured about poor grades?

Visiting Grandparents
The grandfather always told the children about his time in the navy.
* **M↓ Q** of the children were interested in the stories.
* **M↑ Q** of the children were tired of the stories.
[And so][However], they listened carefully until their(her) bedtime came[, however].
Was the grandfather in the navy?

Sitting Exams
The students found that a hectic social life interfered with their studies.
* **M↓ Q** of the students ignored their work.
* **M↑ Q** of the students attended to their work.
[And so][Nevertheless], they failed miserably when the exams were marked[, nevertheless].
Did the students find it easy study and maintain a hectic social life?

In The Pub
Some men were playing on the fruit machine in the pub.
* **M↓ Q** of the men won a lot of money.
* **M↑ Q** of the men lost a lot of money.
[And so][Nevertheless], they celebrated extravagantly until the bar closed[, nevertheless].
Were the men playing on a fruit machine?

At The Protest Match
The animal rights demonstrators were badly treated by the police.
* **M↓ Q** of the demonstrators appeared to condemn the police action.
* **M↑ Q** of the demonstrators appeared satisfied with the police action.
[And so][Nevertheless], they protested vociferously until an apology was given[, nevertheless].
Were the demonstrators well treated by the police?

After The Accident
The police blamed dangerous drivers for a major motorway pile-up.
* **M↓ Q** of the drivers appeared to accept responsibility for the pile-up.
* **M↑ Q** of the drivers refused to accept responsibility for the pile-up.
[And so][Nevertheless], they apologised profusely when the police arrived[, nevertheless].
Did the police blame dangerous drivers for the motorway pile-up?

Catching A Train
Commuters were always in a hurry to catch the London train.
* **M↓ Q** of the commuters thought they were late for the train.
* **M↑ Q** of the commuters thought they were early for the train.
[And so][Nevertheless], they sprinted frantically before the train departed[, nevertheless].
Were the commuters catching the train to Edinburgh?

In The University
A number of university staff were offered early retirement.
* **M↓ Q** of the staff felt like accepting the offer.
* **M↑ Q** of the staff felt like rejecting the offer.
[And so][However], they retired gracefully when the academic year finished[, however].
Did the staff work at the university?

At The School Football Match
Some boys from the local team were playing against a nearby school.
* **M↓ Q** of the boys were hostile to the other team.
* **M↑ Q** of the boys were friendly to the other team.
[And so][However], they argued incessantly until the game was abandoned[, however].
Were the teams playing rugby?

At The Casino
A group of men won a lot of money on the roulette wheel.
* **M↓ Q** of the men wanted to risk their winnings.
* **M↑ Q** of the men wanted to save their winnings.
[And so][However], they gambled recklessly until the money was gone[, however].
Were the men playing roulette at the casino?

A Residents Meeting
Local residents were discussing plans to build a hospital in the area.
* **M↓ Q** of the residents were displeased about the plans.
* **M↑ Q** of the residents were pleased about the plans.
[And so][However], they protested angrily until the plans were dropped[, however].
Did the residents discuss plans to build a prison?

In The Bar
One of the barstaff always made sexist remarks to female customers.
* **M↓ Q** of the women appeared to detest his sense of humour.
* **M↑ Q** of the women appreciated his sense of humour.
[And so][However], they objected vocally when the manager arrived[, however].
Was it one of the barstaff who made sexist remarks to the women?

After The Election
A crowd of people gathered to hear the election result.
* **M↓ Q** of the crowd felt glad about the outcome.
* **M↑ Q** of the crowd felt dismayed by the outcome.
[And so][Nevertheless], they clapped enthusiastically when the winner spoke[, nevertheless].
Did the crowd gather to hear a sporting result?

At Playschool
The children all wanted to play with one of the toys at the same time.
* **M↓ Q** of the children appeared to be aggressive.
* **M↑ Q** of the children appeared to be passive.
[And so][Nevertheless], they fought violently until the toy broke[, nevertheless].
Did the children all want to play with the same toy?

After A Hijacking
The hijackers were trapped in the aircraft and surrounded by police.
* **M↓ Q** of the hijackers wanted to end the siege.
* **M↑ Q** of the hijackers wanted to continue the siege.
[And so][Nevertheless], they surrendered unconditionally before the police attacked[, nevertheless].
Was the aircraft surrounded by the army?

At The Demonstration
The local council wanted to build a new road through the city.
* **M↓ Q** of the local residents seemed to oppose the new road.
* **M↑ Q** of the local residents seemed to approve of the new road.
[And so][Nevertheless], they demonstrated noisily when the workmen arrived[, nevertheless].
Did the local council want to build a new road?

A Student House
A group of students moved into a new house.
* **M↓ Q** of the students usually felt cold in the house.
* **M↑ Q** of the students usually felt warm in the house.
[And so][Nevertheless], they shivered icily until the heating came on[, nevertheless].
Was it a pair of newly-weds who moved into the new house?

Begging In The Street
A tramp was pestering shoppers to give him some more money.
* **M↓ Q** of the shoppers normally agreed to help.
* **M↑ Q** of the shoppers normally refused to help.
[And so][However], they donated generously when the man approached[, however].
Did the tramp ask for money?

After A Robbery
Local residents had suffered a spate of robberies.
* **M↓ Q** of the residents agreed to join the neighbourhood watch.
* **M↑ Q** of the residents refused to join the neighbourhood watch..
[And so][However], they patrolled regularly until the thefts stopped[, however].
Were the residents suffering from vandalism?

In The Folk Club
The folk singer's performance at the club was terrible.
* **M↓ Q** of the folk club members seemed to like the performance.
* **M↑ Q** of the folk club members seemed to dislike the performance.
[And so][However], they clapped loudly at the end[, however].
Was it a folk singer who performed at the club?

After The Last Exam
Some of the students went to the pub after the last exam.
* **M↓ Q** of the students appeared happy about the questions set.
* **M↑ Q** of the students appeared upset about the questions set.
[And so][However], they celebrated enthusiastically at the pub[, however].
Did the students go for a meal after their exam?

A Factory Strike
The workers were holding a ballot in order to decide whether to strike or not.
* **M↓ Q** of the workers seemed in favour of industrial action.
* **M↑ Q** of the workers seemed against industrial action.
[And so][However], they went on strike after the vote[, however].
Did the workers hold a ballot?

Filler items

The following is a list of the 10 filler materials used in Experiment 5.

In Sunday School
The Sunday School class was in full swing.
* **M↓ Q** of the children were happy about going to Sunday School.
* **M↑ Q** of the children were unhappy about going to Sunday School.
And so, they complained to their parents when they returned home.
Were the children attending Sunday School?

At School
It was the first day back at school after the holidays.
* **M↓ Q** of the children wanted to go back to school.
* **M↑ Q** of the children didn't want to go back to school.
And so, they complained miserably when the first bell rang .
Was it the last day of term?

At The Evening Class
The Computer Skills evening class was about to begin.
* **M↓ Q** of the people had used a computer before.
* **M↑ Q** of the people hadn't used a computer before.
And so, they nervously waited for the class to start.
Were the students attending an evening class?

In The Park
The sun shone as the students lazed in the park.
* **M↓ Q** of the students wanted to go to their next lecture.
* **M↑ Q** of the students didn't want to go to their next lecture.
And so, they happily lay in the sun until evening.
Were the students in a pub?

UFO Spotting
The UFO spotters gazed up at the night sky.
* **M↓ Q** of the crowd felt that they wouldn't see a UFO that evening.
* **M↑ Q** of the crowd felt that they would see a UFO that evening.
And so, they watched expectantly throughout the night.
Were the crowd trainspotting?

A Night Out
A group of students were drinking in the pub.
* **M↓ Q** of the students wanted to stay sober.
* **M↑ Q** of the students wanted to get drunk.
And so, the drank enthusiastically all night.
Were the students at a concert?

At The Amusement Park
Some children had been taken for a day out at the amusement park.
* **M↓ Q** of the children were having a boring time.
* **M↑ Q** of the children were having an exciting time.
And so, they played happily all day.
Were the children at an amusement park?

Writing Essays
The students had been given a series of essays to write for the course.
* **M↓ Q** of the students were good at writing essays.
* **M↑ Q** of the students were poor at writing essays.
And so, they worked hard until their writing improved.
Were the students studying for an exam?

Driving To Work
A group of motorists got stuck in a traffic jam on their way to work.
* **M↓ Q** of the motorists were pleased about the delay.
* **M↑ Q** of the motorists were angry about the delay
And so, they waited impatiently until the traffic started moving again.
Were the motorists going to work?

A New Term
A group of teachers were preparing for the new term.
* **M↓ Q** of the teachers were unmotivated by their work.
* **M↑ Q** of the teachers were motivated by their work
And so, they taught enthusiastically throughout the year.
Was it a group of teachers who were preparing for the new term?

Appendix 3:

Experimental materials for Experiment 6

The following list relates the experimental materials used in Experiment 6. For the sake of clarity, materials from the *simple mismatch* and *adversative* conditions appear in a separate list from the *focus match* condition. In these materials, the symbol "*" denotes the between subjects (*quantifier*) variants and "[ ]" denotes adversative variants of the *anaphoric reference sentence.*

Materials from the *simple mismatch* and *adversative* conditions:

At The Boy-scout Camp
It was late before the boy-scouts had finished setting-up camp.
* Many of the boys felt wide awake after the work.
* Not many of the boys felt very tired after the work.
* Simon felt wide awake after the work.
[Nevertheless,]They(*He*) slept soundly[, nevertheless,]until morning arrived.
They(*He*) awoke feeling refreshed and ready for the day ahead.
Was the camp set-up by girl-guides?

In A Retirement Home
All of the residents were in the TV room during Brookside.
* Not many of the residents were excited by the plot.
* Many of the residents were bored by the plot.
* Emma was bored by the plot.
[Nevertheless,]They(She) watched intently[, nevertheless,] when the programme began.
They(She) began chatting when the end credits rolled.
Were the residents watching Brookside?

Caught Shop-Lifting
A group of school-boys were suspected of stealing from the local shop.
* Not many of the boys appeared guilty of the crime.
* Many of the boys appeared innocent of the crime.
* Henry appeared innocent of the crime.
[Nevertheless,]They(*He*) confessed immediately[, nevertheless,] when the police arrived.
They were(He was) arrested and taken to the police station.
Were the boys suspected of vandalism?

At The Theatre
The local theatre company put on a controversial work.
* Not many of the audience felt pleased by the performance.
* Many of the audience felt disgusted by the performance.
* Sheila felt disgusted by the performance.
[Nevertheless,]They(She) applauded appreciatively[, nevertheless,] when the play finished.
Their(Her) arms were tired by the time the applause finished.
Did the theatre company put on a controversial work?

At A Conference
The conference delegates used the university cafeteria for lunch.
* Not many of the delegates appeared disappointed by the food.
* Many of the delegates appeared impressed by the food.
* Sandra appeared impressed by the food
[However,]They(She) complained bitterly[, however,] when a manager arrived.
Their(Her) harsh words clearly bothered the staff.

Did the conference delegates go to the pub for lunch?

Keeping Fit

The security guards were almost all unfit and overweight.

* Not many of the men were enthusiastic about improving their health.
* Many of the men were apathetic about improving their health.
* Roger was apathetic about improving his health.

[However,]They(He) exercised vigorously[, however,] before their(his) shift began.
They(He) felt exhausted by the time they(he) had finished.
Were the security guards unfit and overweight?

In A Hospital

Some medical students had the chance to participate in a drugs trial.

* Not many of the students appeared reluctant to participate.
* Many of the students appeared keen to participate.
* Gillian appeared keen to participate.

[However,]They(She) refused pointedly[, however,] when the medical authorities asked.
Their(Her) refusal was accepted without question.
Were the students studying law?

Going To A Party

Some of the school-boys asked their parents for beer to take to the party.

* Not many of the parents felt happy about underage drinking.
* Many of the parents felt upset about underage drinking.
* Andrew, one of the parents, felt upset about underage drinking.

[However,]They (He) agreed willingly[, however,] when the boys asked.
They(He) refused to take any money for the beer.
Did the schoolboys ask their parents for beer?

In The Classroom

The teacher lectured the entire class about behaving properly.

* Not many of the children were scolded by the teacher.
* Many of the children were praised by the teacher.
* Sally was praised by the teacher.

[Nevertheless,]They(She) wept tearfully[, nevertheless,] until their parents arrived.
Their(Her) tears concerned the teacher.
Was the class lectured about poor grades?

Visiting Grandparents

The grandfather always told the children about his time in the navy.

* Not many of the children were interested in the stories.
* Many of the children were tired of the stories.
* Sharon was tired of the stories.

[Nevertheless,]They(She) listened carefully[, nevertheless,] until their(her) bedtime came.
Their(Her) attentiveness was evident to their(her) grandfather.
Was the grandfather in the navy?

Sitting Exams

The students found that a hectic social life interfered with their studies.

* Not many of the students ignored their work.
* Many of the students attended to their work.
* Kim attended to her work.

[Nevertheless,]They(She) failed miserably[, nevertheless,] when the exams were marked.
They were(she was) very disappointed by the results.
Did the students find it easy study and maintain a hectic social life?

In The Pub
Some men were playing on the fruit machine in the pub.
* Not many of the men won a lot of money.
* Many of the men lost a lot of money.
* Tom lost a lot of money.
[Nevertheless,]They(He) celebrated extravagantly[, nevertheless,] until the bar closed.
Their(His) good humour was remarked upon by the barstaff.
Were the men playing on a fruit machine?

At The Protest Match
The animal rights demonstrators were badly treated by the police.
* Not many of the demonstrators appeared to condemn the police action.
* Many of the demonstrators appeared satisfied with the police action.
* Allison appeared satisfied with the police action.
[However,]They(She) protested vociferously[, however,] until an apology was given.
Their(Her) protest clearly had the intended effect.
Were the demonstrators well treated by the police?

After The Accident
The police blamed dangerous drivers for a major motorway pile-up.
* Not many of the drivers appeared to accept responsibility for the pile-up.
* Many of the drivers refused to accept responsibility for the pile-up.
* Jeremy refused to accept responsibility for the pile-up.
[However,]They(He) apologised profusely[, however,] when the police arrived.
They were(He was) cautioned about their(his) driving by the police.
Did the police blame dangerous drivers for the motorway pile-up?

Catching A Train
Commuters were always in a hurry to catch the London train.
* Not many of the commuters thought they were late for the train.
* Many of the commuters thought they were early for the train.
* Michelle thought she was early for the train.
[However,]They(She) sprinted frantically[, however,] before the train departed.
They were(She was) breathless by the time they(she) reached the train.
Were the commuters catching the train to Edinburgh?

In The University
A number of university staff were offered early retirement.
* Not many of the staff felt like accepting the offer.
* Many of the staff felt like rejecting the offer.
* Robert felt like rejecting the offer.
[However,]They(He) retired gracefully[, however,] when the academic year finished.
Their(His) friends at the university were sad to see them(him) go.
Did the staff work at the university?

At The School Football Match
Some boys from the local team were playing against a nearby school.
* Not many of the boys were hostile to the other team.
* Many of the boys were friendly to the other team.
* Collin was friendly to the other team.
[Nevertheless,]They(He) argued incessantly[, nevertheless,] until the game was abandoned.
They were(He was) disappointed that the game was stopped.
Were the teams playing rugby?

At The Casino
A group of men won a lot of money on the roulette wheel.
* Not many of the men wanted to risk their winnings.
* Many of the men wanted to save their winnings.
* Kenny wanted to save his winnings.
[Nevertheless,]They(He) gambled recklessly[, nevertheless,] until the(his) money was gone.
They(He) decided to be more careful the next time.
Were the men playing roulette at the casino?

A Residents Meeting
Local residents were discussing plans to build a hospital in the area.
* Not many of the residents were displeased about the plans.
* Many of the residents were pleased about the plans.
* Shona was pleased about the plans.
[Nevertheless,]They(She) protested angrily[, nevertheless,] until the plans were dropped.
They were(She was) happy that their(her) protest was successful.
Did the residents discuss plans to build a prison?

In The Bar
One of the barstaff always made sexist remarks to female customers.
* Not many of the women appeared to detest his sense of humour.
* Many of the women appreciated his sense of humour.
* Bellinda appreciated his sense of humour.
[Nevertheless,]They(She) objected vocally[, nevertheless,] when the manager arrived.
Their(Her) annoyance was obvious to the manager.
Was it one of the barstaff who made sexist remarks to the women?

After The Election
A crowd of people gathered to hear the election result.
* Not many of the crowd felt glad about the outcome.
* Many of the crowd felt dismayed by the outcome.
* Helen felt dismayed by the outcome.
[However,]They(She) clapped enthusiastically[, however,] when the winner spoke.
Their(Her) response clearly pleased the winner.
Did the crowd gather to hear a sporting result?

At Playschool
The children all wanted to play with one of the toys at the same time.
* Not many of the children appeared to be aggressive.
* Many of the children appeared to be passive.
* Stephen appeared to be passive.
[However,]They(He) fought violently[, however,] until the toy broke.
They(He) decided that they(he) would behave better in the future.
Did the children all want to play with the same toy?

After A Hijacking
The hijackers were trapped in the aircraft and surrounded by police.
* Not many of the hijackers wanted to end the siege.
* Many of the hijackers wanted to continue the siege.
* Fred, one of the hijackers, wanted to continue the siege.
[However,]They(He) surrendered unconditionally[, however,] before the police attacked.
They were(He was) quickly bundled into a police van and driven off.
Was the aircraft surrounded by the army?

At The Demonstration
The local council wanted to build a new road through the city.
* Not many of the local residents seemed to oppose the new road.
* Many of the local residents seemed to approve of the new road.
* Mary seemed to approve of the new road.
[However,]They(She) demonstrated noisily[, however,] when the workmen arrived.
Their(Her) actions delayed the beginning of the work.
Did the local council want to build a new road?

A Student House
A group of students moved into a new house.
* Not many of the students usually felt cold in the house.
* Many of the students usually felt warm in the house.
* Gemma usually felt warm in the house.
[Nevertheless,]They(She) shivered icily[, nevertheless,] until the heating came on.
They(She) felt happier once they were(she was) warm.
Was it a pair of newly-weds who moved into the new house?

Begging In The Street
A tramp was pestering shoppers to give him some more money.
* Not many of the shoppers normally agreed to help.
* Many of the shoppers normally refused to help.
* Brian normally refused to help.
[Nevertheless,]They(He) donated generously[, nevertheless,] when the man approached.
Their(His) money was gratefully accepted by the man.
Did the tramp ask for money?

After A Robbery
Local residents had suffered a spate of robberies.
* Not many of the residents agreed to join the neighbourhood watch.
* Many of the residents refused to join the neighbourhood watch.
* Jill refused to join the neighbourhood watch.
[Nevertheless,]They(She) patrolled regularly[, nevertheless,] until the thefts stopped.
They were(She was) pleased by the success of their(her) action.
Were the residents suffering from vandalism?

In The Folk Club
The folk singer's performance at the club was terrible.
* Not many of the folk club members seemed to like the performance.
* Many of the folk club members seemed to dislike the performance.
* Tony seemed to dislike the performance.
[Nevertheless,]They(He) clapped loudly[, nevertheless,] at the end.
They(He) stayed on for a drink at the club's bar.
Was it a folk singer who performed at the club?

After The Last Exam
Some of the students went to the pub after the last exam.
* Not many of the students appeared happy about the questions set.
* Many of the students appeared upset about the questions set.
* Mary appeared upset about the questions set.
[However,]They(She) celebrated enthusiastically[, however,] at the pub.
They(She) hoped that they(She) wouldn't feel too bad the next day.
Did the students go for a meal after their exam?

A Factory Strike
The workers were holding a ballot in order to decide whether to strike or not.
* Not many of the workers seemed in favour of industrial action.
* Many of the workers seemed against industrial action.
* Henry seemed against industrial action.
[However,]They(He) went on strike[, however,] after the vote.
They(He) hoped that the strike would be effective.
Did the workers hold a ballot?

At School
It was the first day back at school after the holidays.
* Not many of the children didn't want to go back to school.
* Many of the children wanted to go back to school.
* Margaret wanted to go back to school.
[However,]They(She) complained miserably[, however,] when the first bell rang.
They(She) walked dejectedly into the classroom.
Was it the last day of term?

At the Evening Class
The Computer Skills evening class was about to begin.
* Not many of the people hadn't used a computer before.
* Many of the people had used a computer before.
* Arthur had used a computer before.
[However,]They(He) nervously waited[, however,] for the class to start.
They(He) hoped that the course wouldn't be too hard.
Were the students attending an evening class?

Materials from the *focus match* condition:

At The Boy-scout Camp
It was late before the boy-scouts had finished setting-up camp.
* Many of the boys felt very tired after the work.
* Not many of the boys felt wide awake after the work.
* Simon felt very tired after the work.
They(*He*) slept soundly until morning arrived.
They(*He*) awoke feeling refreshed and ready for the day ahead.
Was the camp set-up by girl-guides?

In A Retirement Home
All of the residents were in the TV room during Brookside.
* Not many of the residents were bored by the plot.
* Many of the residents were excited by the plot.
* Emma was excited by the plot.
They(She) watched intently when the programme began.
They(She) began chatting when the end credits rolled.
Were the residents watching Brookside?

Caught Shop-Lifting
A group of school-boys were suspected of stealing from the local shop.
* Not many of the boys appeared innocent of the crime.
* Many of the boys appeared guilty of the crime.
* Henry appeared guilty of the crime.
They(*He*) confessed immediately when the police arrived.
They were(He was) arrested and taken to the police station.
Were the boys suspected of vandalism?

At The Theatre
The local theatre company put on a controversial work.
* Not many of the audience felt disgusted by the performance.
* Many of the audience felt pleased by the performance.
* Sheila felt pleased by the performance.
They(She) applauded appreciatively when the play finished.
Their(Her) arms were tired by the time the applause finished.
Did the theatre company put on a controversial work?

At A Conference
The conference delegates used the university cafeteria for lunch.
* Not many of the delegates appeared impressed by the food.
* Many of the delegates appeared disappointed by the food.
* Sandra appeared disappointed by the food
They(She) complained bitterly when a manager arrived.
Their(Her) harsh words clearly bothered the staff.
Did the conference delegates go to the pub for lunch?

Keeping Fit
The security guards were almost all unfit and overweight.
* Not many of the men were apathetic about improving their health.
* Many of the men were enthusiastic about improving their health.
* Roger was enthusiastic about improving his health.
They(He) exercised vigorously before their(his) shift began.
They(He) felt exhausted by the time they(he) had finished.
Were the security guards unfit and overweight?

In A Hospital
Some medical students had the chance to participate in a drugs trial.
* Not many of the students appeared keen to participate.
* Many of the students appeared reluctant to participate.
* Gillian appeared reluctant to participate.
They(She) refused pointedly when the medical authorities asked.
Their(Her) refusal was accepted without question.
Were the students studying law?

Going To A Party
Some of the school-boys asked their parents for beer to take to the party.
* Not many of the parents felt upset about underage drinking.
* Many of the parents felt happy about underage drinking.
* Andrew, one of the parents, felt happy about underage drinking.
They (He) agreed willingly when the boys asked.
They(He) refused to take any money for the beer.
Did the schoolboys ask their parents for beer?

In The Classroom
The teacher lectured the entire class about behaving properly.
* Not many of the children were praised by the teacher.
* Many of the children were scolded by the teacher.
* Sally was scolded by the teacher.
They(She) wept tearfully until their parents arrived.
Their(Her) tears concerned the teacher.
Was the class lectured about poor grades?

Visiting Grandparents
The grandfather always told the children about his time in the navy.
* Not many of the children were tired of the stories.
* Many of the children were interested in the stories.
* Sharon was interested in the stories.
They(She) listened carefully until their(her) bedtime came.
Their(Her) attentiveness was evident to their(her) grandfather.
Was the grandfather in the navy?

Sitting Exams
The students found that a hectic social life interfered with their studies.
* Not many of the students attended to their work.
* Many of the students ignored their work.
* Kim ignored her work.
They(She) failed miserably when the exams were marked.
They were(she was) very disappointed by the results.
Did the students find it easy study and maintain a hectic social life?

In The Pub
Some men were playing on the fruit machine in the pub.
* Not many of the men lost a lot of money.
* Many of the men won a lot of money.
* Tom won a lot of money.
They(He) celebrated extravagantly until the bar closed.
Their(His) good humour was remarked upon by the barstaff.
Were the men playing on a fruit machine?

At The Protest Match
The animal rights demonstrators were badly treated by the police.
* Not many of the demonstrators appeared satisfied with the police action.
* Many of the demonstrators appeared to condemn the police action.
* Allison appeared to condemn the police action.
They(She) protested vociferously until an apology was given.
Their(Her) protest clearly had the intended effect.
Were the demonstrators well treated by the police?

After The Accident
The police blamed dangerous drivers for a major motorway pile-up.
* Not many of the drivers refused to accept responsibility for the pile-up.
* Many of the drivers appeared to accept responsibility for the pile-up.
* Jeremy appeared to accept responsibility for the pile-up.
They(He) apologised profusely when the police arrived.
They were(He was) cautioned about their(his) driving by the police.
Did the police blame dangerous drivers for the motorway pile-up?

Catching A Train
Commuters were always in a hurry to catch the London train.
* Not many of the commuters thought they were early for the train.
* Many of the commuters thought they were late for the train.
* Michelle thought she was late for the train.
They(She) sprinted frantically before the train departed.
They were(She was) breathless by the time they(she) reached the train.
Were the commuters catching the train to Edinburgh?

In The University
A number of university staff were offered early retirement.
* Not many of the staff felt like rejecting the offer.
* Many of the staff felt like accepting the offer.
* Robert felt like accepting the offer.
They(He) retired gracefully when the academic year finished.
Their(His) friends at the university were sad to see them(him) go.
Did the staff work at the university?

At The School Football Match
Some boys from the local team were playing against a nearby school.
* Not many of the boys were friendly to the other team.
* Many of the boys were hostile to the other team.
* Collin was hostile to the other team.
They(He) argued incessantly until the game was abandoned.
They were(He was) disappointed that the game was stopped.
Were the teams playing rugby?

At The Casino
A group of men won a lot of money on the roulette wheel.
* Not many of the men wanted to save their winnings.
* Many of the men wanted to risk their winnings.
* Kenny wanted to risk his winnings.
They(He) gambled recklessly until the(his) money was gone.
They(He) decided to be more careful the next time.
Were the men playing roulette at the casino?

A Residents Meeting
Local residents were discussing plans to build a hospital in the area.
* Not many of the residents were pleased about the plans.
* Many of the residents were displeased about the plans.
* Shona was displeased about the plans.
They(She) protested angrily until the plans were dropped.
They were(She was) happy that their(her) protest was successful.
Did the residents discuss plans to build a prison?

In The Bar
One of the barstaff always made sexist remarks to female customers.
* Not many of the women appreciated his sense of humour.
* Many of the women appeared to detest his sense of humour.
* Bellinda appeared to detest his sense of humour.
They(She) objected vocally when the manager arrived.
Their(Her) annoyance was obvious to the manager.
Was it one of the barstaff who made sexist remarks to the women?

After The Election
A crowd of people gathered to hear the election result.
* Not many of the crowd felt dismayed by the outcome.
* Many of the crowd felt glad about the outcome.
* Helen felt glad about the outcome.
They(She) clapped enthusiastically when the winner spoke.
Their(Her) response clearly pleased the winner.
Did the crowd gather to hear a sporting result?

At Playschool
The children all wanted to play with one of the toys at the same time.
* Not many of the children appeared to be passive.
* Many of the children appeared to be aggressive.
* Stephen appeared to be aggressive.
They(He) fought violently until the toy broke.
They(He) decided that they(he) would behave better in the future.
Did the children all want to play with the same toy?

After A Hijacking
The hijackers were trapped in the aircraft and surrounded by police.
* Not many of the hijackers wanted to continue the siege.
* Many of the hijackers wanted to end the siege.
* Fred, one of the hijackers, wanted to end the siege.
They(He) surrendered unconditionally before the police attacked.
They were(He was) quickly bundled into a police van and driven off.
Was the aircraft surrounded by the army?

At The Demonstration
The local council wanted to build a new road through the city.
* Not many of the local residents seemed to approve of the new road.
* Many of the local residents seemed to oppose the new road.
* Mary seemed to oppose the new road.
They(She) demonstrated noisily when the workmen arrived.
Their(Her) actions delayed the beginning of the work.
Did the local council want to build a new road?

A Student House
A group of students moved into a new house.
* Not many of the students usually felt warm in the house.
* Many of the students usually felt cold in the house.
* Gemma usually felt cold in the house.
They(She) shivered icily until the heating came on.
They(She) felt happier once they were(she was) warm.
Was it a pair of newly-weds who moved into the new house?

Begging In The Street
A tramp was pestering shoppers to give him some more money.
* Not many of the shoppers normally refused to help.
* Many of the shoppers normally agreed to help.
* Brian normally agreed to help.
They(He) donated generously when the man approached.
Their(His) money was gratefully accepted by the man.
Did the tramp ask for money?

After A Robbery
Local residents had suffered a spate of robberies.
* Not many of the residents refused to join the neighbourhood watch.
* Many of the residents agreed to join the neighbourhood watch.
* Jill agreed to join the neighbourhood watch.
They(She) patrolled regularly until the thefts stopped.
They were(She was) pleased by the success of their(her) action.
Were the residents suffering from vandalism?

In The Folk Club
The folk singer's performance at the club was terrible.
* Not many of the folk club members seemed to dislike the performance.
* Many of the folk club members seemed to like the performance.
* Tony seemed to like the performance.
They(He) clapped loudly at the end.
They(He) stayed on for a drink at the club's bar.
Was it a folk singer who performed at the club?

After The Last Exam
Some of the students went to the pub after the last exam.
* Not many of the students appeared upset about the questions set.
* Many of the students appeared happy about the questions set.
* Mary appeared happy about the questions set.
They(She) celebrated enthusiastically at the pub.
They(She) hoped that they(She) wouldn't feel too bad the next day.
Did the students go for a meal after their exam?

A Factory Strike
The workers were holding a ballot in order to decide whether to strike or not.
* Not many of the workers seemed against industrial action.
* Many of the workers seemed in favour of industrial action.
* Henry seemed in favour of industrial action.
They(He) went on strike after the vote.
They(He) hoped that the strike would be effective.
Did the workers hold a ballot?

At School
It was the first day back at school after the holidays.
* Not many of the children wanted to go back to school.
* Many of the children didn't want to go back to school.
* Margaret didn't want to go back to school.
They(She) complained miserably when the first bell rang.
They(She) walked dejectedly into the classroom.
Was it the last day of term?

At the Evening Class
The Computer Skills evening class was about to begin.
* Not many of the people had used a computer before.
* Many of the people hadn't used a computer before.
* Arthur hadn't used a computer before.
They(He) nervously waited for the class to start.
They(He) hoped that the course wouldn't be too hard.
Were the students attending an evening class?

Appendix 4:

Experimental materials for Experiment 7

The following list relates the 32 experimental materials used in Experiment 7. In these, slashes (/) denote alternatives and hashes (#) denote region divisions.

At The Boy-scout Camp

It was late before the boy-scouts had finished setting-up camp. # A few/Few of the boys felt very tired/wide awake after the work, so # they slept soundly # until morning arrived.

In a retirement home

All of the residents were usually in the TV room during Brookside. # A few/Few of the residents were excited/bored with the plot, so # they watched intently # when the programme began. Was Coronation Street on the TV?

Caught shop-lifting

A group of school-boys were suspected of stealing from the local shop. # A few/Few of the boys were innocent/guilty of the crime, so # they confessed immediately # when the police arrived.

At the theatre

The local theatre company put on a controversial work. # A few/Few of the audience were pleased/disgusted by the performance, so # they applauded appreciatively # when the play finished.

At a conference

The conference delegates used the university cafeteria for lunch. # A few/Few of the delegates were impressed/disappointed with the food, so # they complained bitterly # when a manager arrived. Did the conference delegates use the university cafeteria?

Keeping fit

The security guards were almost all unfit and overweight. # A few/Few of the men were enthusiastic/apathetic about improving their health, so # they exercised vigorously # before the shift started.

In a hospital

Some medical students had the chance to participate in a drugs trial. # A few/Few of the students were keen/reluctant to take part, so # they refused pointedly # when the medical authorities asked.

Going to a party

Some of the school-boys asked their parents for beer to take to the party. # A few/Few of the parents allowed/forbade underage drinking, so # they agreed willingly # when the boys asked.

In the classroom

The teacher lectured the entire class about behaving properly. # A few/Few of the children were praised/scolded by the teacher, so # they wept tearfully # until their parents arrived.

Visiting grandparents

The grandfather always told the children about his time in the navy. # A few/Few of the children were interested in/tired of the stories, so # they listened carefully # until their bedtime came. Was the grandfather in the navy?

Sitting exams

The students found that a hectic social life interfered with their studies. # A few/Few of the students attended to/ignored their work, so # they failed miserably # when the exams were marked.

In the pub

Some men were playing on the fruit machine in the pub. # A few/Few of the men won/lost a lot of money, so # they celebrated extravagantly # until the bar closed.

Catching a bus

As usual the bus was busy with commuters during the rush hour. # A few/Few of the commuters managed/failed to get a seat, so # they stood awkwardly # until the journey was over. Did the commuters travel by train?

Catching flu

The school was warned that a lot of children might fall prey to a flu epidemic. # A few/Few of the children caught/avoided the virus, so # they sneezed constantly # until the infection had passed.

At the protest match

The animal rights demonstrators were badly treated by the police. # A few/Few of the demonstrators supported/condemned the police action, so # they protested vociferously # until an apology was given.

After the accident

The police blamed dangerous drivers for a major motorway pile-up. # A few/Few of the drivers admitted/denied responsibility, so # they apologised profusely # when the police arrived.

Catching a train

Commuters were always in a hurry to catch the London train. # A few/Few of the commuters were early/late for the train, so # they sprinted frantically # before the train departed.

In the university

A number of university staff were offered early retirement. # A few/Few of the staff accepted/rejected the offer, so # they retired gracefully # when the academic year finished.

At the school football match

Some boys from the local team were playing against a nearby school. # A few/Few of the boys were friendly/hostile to the other team, so # they argued incessantly # until the game was abandoned.

After a shipwreck

Most of the passengers dived overboard when the ship sank. # A few/Few of the passengers managed/struggled to stay afloat, so # they waited calmly # until the rescue services arrived.
Did most of the passengers dive overboard?

At the casino

A group of men won a lot of money on the roulette wheel. # A few/Few of the men were careful/careless with their winnings, so # they gambled recklessly # until the money was gone.
Were the men playing poker?

In parliament

MPs were scheduled to debate increased taxation for gas and electricity. # A few/Few of the MPs were present/absent from the debate, so # they voted unanimously # after the motion was debated.

In the bar

One of the barstaff always made sexist remarks to female customers. # A few/Few of the women appreciated/detested his sense of humour, so # they objected vocally # when the manager arrived.

After the election

A crowd of people gathered to hear the election result. # A few/Few of the crowd felt glad/sad about the outcome, so # they clapped enthusiastically # when the winner spoke.

At playschool

The children all wanted to play with one of the toys at the same time. # A few/Few of the children were passive/aggressive by nature, so # they fought violently # until the toy broke

After a hijacking
The hijackers were trapped in the aircraft and surrounded by police. # A few/Few of the hijackers wanted to continue/end the siege, so # they surrendered unconditionally # before the police attacked.

At the demonstration

The local council wanted to build a new road through the city. # A few/Few of the local residents approved/opposed the new road, so # they demonstrated noisily when the workmen arrived.
Did the local council want to build a road?

At the trial

Some Irishmen living in London were accused of terrorist activities. # A few/Few of the Irishmen were acquitted/convicted at the trial, so # they rejoiced triumphantly # when they were released.
Were the Irishmen living in Belfast?

A student house

A group of students moved into a new house.
A few/Few of the students felt warn/cold in the house, so # they shivered icily # until the heating came on.

Begging in the street

A tramp was pestering shoppers to give him some money. # A few/Few of the shoppers offered/refused to help, so # they donated generously # when the man approached
Was the money being requested by a charity collector?

On holiday abroad

A group of tourists drank untreated tap water in the hotel.
A few/Few of the tourists suffered/avoided a bout of food poisoning,
so # they vomited copiously # until the doctor arrived.

After A Robbery
Local residents had suffered a spate of robberies.
A few/Few of the residents agreed/declined to join the neighbourhood
watch, so # they patrolled regularly # until the thefts stopped.